

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleotide search, using frame_plus.p2n model

Run on: July 11, 2006, 19:46:18 ; Search time 1356.71 Seconds
(without alignments)
4383.493 Million cell updates/sec

Title: US-10-718-311-15
Perfect score: 305
Sequence: 1 MASSVISAATRSNTVQA.....LDITSIASNGRVSQMWH 62

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/ABSSWEB spool/US10718311/runat_11072006_111520_5171/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes03p
-USER=US10718311 @CGN 1 1 8663 @runat_11072006_111520_5171 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb_env.*
- 2: gb_pat.*
- 3: gb_ph.*
- 4: gb_pl.*
- 5: gb_pr.*
- 6: gb_ro.*
- 7: gb_sts.*
- 8: gb_sy.*
- 9: gb_un.*
- 10: gb_vi.*
- 11: gb_ov.*
- 12: gb_htg.*
- 13: gb_in.*
- 14: gb_om.*
- 15: gb_ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	305	100.0	684	2 AR428736 Sequence
2	305	100.0	684	2 AR455412 Sequence
3	305	100.0	684	2 AX329368 Sequence

4	305	100.0	684	2	AX382258	Sequence
5	297	97.4	599	4	TOMRBCSB	M13543 Tomato (L.e
6	297	97.4	742	4	TOMRBCSE	M15236 Tomato RuBP
7	297	97.4	1097	4	BT013023	BT013023 Lycopersi
8	292	95.7	2776	4	LERBCS2	X05983 Tomato rbcS
9	269	88.2	204	2	AR528363	AR528363 Sequence
10	269	88.2	499	2	AR205973	AR205973 Sequence
11	269	88.2	796	4	NSRUB1	X01722 Nicotiana s
12	266	87.2	806	4	AY220079	AY220079 Nicotiana
13	262	85.9	22574	8	CVE18556	M18556 Cloning vec
14	258	84.6	979	4	TORUBPA	M32419 Tobacco rib
15	258	84.6	2293	4	NPRBCS8B	X13711 Nicotiana p
16	258	84.6	2293	4	TORBCS8B	M36685 N.plumbagin
17	258	84.6	2362	4	NTRUBS	X02353 Tobacco gen
18	257.5	84.4	546	4	POTRBCS	J03613 Potato (S.t
19	256.5	84.1	3323	4	STRBCS1	X69759 S.tuberosum
20	256	83.9	723	4	SLARBCS	L26605 Stellaria l
21	254	83.3	177	2	I24839	I24839 Sequence 8
22	254	83.3	177	2	I40228	I40228 Sequence 6
23	254	83.3	177	2	I72655	I72655 Sequence 12
24	254	83.3	225	8	ARPRECA2	X05923 Hybrid rbc-
25	254	83.3	1054	4	LERBCS3B	X05985 Tomato rbcS
26	254	83.3	4124	4	TOMRBCO	D11112 Tomato ribu
27	253.5	83.1	778	4	TOMRBCSD	M15235 Tomato RuBP
28	253	83.0	1027	4	TOMRBCSC	M13544 Tomato (L.e
29	253	83.0	1341	4	LERBCS3C	X05986 Tomato rbcS
30	253	83.0	1520	4	LERBCS3A	X05984 Tomato rbcS
31	251	82.3	1261	4	BT012936	BT012936 Lycopersi
32	249	81.6	2242	4	PERBCS08	X03820 Petunia x h
33	248	81.3	1386	4	STRBCS2C	X69762 S.tuberosum
34	248	81.3	1629	4	STRBCS2	X69760 S.tuberosum
35	247	81.0	1314	4	PERBCS11	X03821 Petunia x h
36	246.5	80.8	1598	4	STRBCS3	X69763 S.tuberosum
37	246	80.7	1703	4	STRBCS2B	X69761 S.tuberosum
38	242.5	79.5	1032	4	TOMRBCSA	M13542 Tomato (L.e
39	242.5	79.5	1454	4	LERBCS1	X05982 Tomato rbcS
40	228.5	74.9	1337	4	TORUBPB	M32420 Tobacco rib
41	228	74.8	801	4	AF065615	AF065615 Capsicum
42	225	73.8	546	2	CQ805080	CQ805080 Sequence
43	225	73.8	546	4	AY098970	AY098970 Arabidops
44	225	73.8	577	4	AY051025	AY051025 Arabidops
45	225	73.8	595	4	AY064686	AY064686 Arabidops

ALIGNMENTS

RESULT 1
AR428736
LOCUS AR428736 684 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 16 from patent US 6642036.
ACCESSION AR428736
VERSION AR428736.1 GI:40188466
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 684)
AUTHORS Flint, D., Meyer, K. and Viitanen, P.
TITLE Sinapoylglucose:malate sinapoyltransferase form malate conjugates from benzoic acid glucosides
JOURNAL Patent: US 6642036-A 16 04-NOV-2003;
E. I. Du Pont de Nemours and Company; Wilmington;

WOX;

FEATURES

source

Location/Qualifiers

1..684

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 4.56e-29 Length: 684

Score: 305.00 Matches: 62

Percent Similarity: 100.0% Conservative: 0

```

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AR428736 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

QY 61 TrpHis 62
DB 181 TGGCAT 186

RESULT 2
AR455412
LOCUS AX329368 684 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 7 from patent US 6683231.
ACCESSION AR455412
VERSION AR455412.1 GI:42689960
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
source
1. .684
/mol_type="unknown"
/organism="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservat: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AR455412 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

QY 61 TrpHis 62
DB 181 TGGCAT 186

RESULT 3
AX329368
LOCUS AX329368 684 bp DNA linear PAT 08-JAN-2002
DEFINITION Sequence 7 from Patent WO0194607.

```

```

ACCESSION AX329368
VERSION AX329368.1 GI:18102367
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
E.I. Dupont De Nemours (US)
FEATURES
source
1. .684
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic CPL"

ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservat: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AX329368 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTTCACCTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

QY 61 TrpHis 62
DB 181 TGGCAT 186

RESULT 4
AX382258
LOCUS AX382258 684 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 16 from Patent WO0204653.
ACCESSION AX382258
VERSION AX382258.1 GI:19577037
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
E.I. DUPONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1. .684
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="open reading frame of the chloroplast-targeted CPL fusion protein"

ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 684
Score: 305.00 Matches: 62

```

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x AX382258 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 1 ATGGCTCTCTGTCATTTCTTCAGCAGCTGTTGCCACACGACGCAATGTTACACAAGCT 60
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrLysLys 40
 |||||
 Db 61 AGCATGGTTCACCTTCTACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
 |||||

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 121 CAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
 |||||

QY 61 TrpHis 62
 |||||
 Db 181 TGGCAT 186

RESULT 5
 TOMRBCSB
 LOCUS 599 bp mRNA linear PLN 27-APR-1993
 DEFINITION Tomato (L.esculentum) ribulose-1,5-bisphosphate
 carboxylase/oxygenase (RBCS) small subunit mRNA, clone Rbcs-2A.
 ACCESSION M13543
 VERSION M13543.1 GI:170497
 KEYWORDS ribulose 1,5-bisphosphate carboxylase/oxygenase.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
 AUTHORS Pichersky, E., Bernatzky, R., Tanksley, S.D. and Cashmore, A.R.
 TITLE Evidence for selection as a mechanism in the concerted evolution of
 Lycopersicon esculentum (tomato) genes encoding the small subunit
 of ribulose-1,5-bisphosphate carboxylase/oxygenase
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (11), 3880-3884 (1986)
 PUBMED 3012537
 COMMENT Original source text: Tomato (L.esculentum) cDNA to mRNA, clone
 Rbcs-2A.
 Clean copy sequence for [1] kindly provided by E.Pichersky,
 31-OCT-1986

FEATURES
 source Location/Qualifiers
 1..599
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 15..557
 /note="ribulose-1,5-bisphosphate carboxylase/ oxygenase
 small subunit (EC 4.1.1.39)"
 /codon_start=1
 /protein_id="AAA34189.1"
 /db_xref="GI:170498"
 /translation="MASSVITSSAAVATRSNVTQASVAPFTGLKSSATPVTKKQNLID
 ITSISANGRVSCMVQWPPINMKKYETLSVLPDLSEQLLSEVYLLKNGVPCLEFE
 TEGFVYRENNKSGFYDGRYTWKMLPMFGCTDTAQVLAEOEAKKAYPAQWVRIG
 FDNVRQVCISFIAYKPEGY"

CDS
 ORIGIN Chromosome 3, Rbcs-2 locus.
 Alignment Scores:
 Pred. No.: 3,99e-28 Length: 599
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSB (1-599)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 15 ATGGCTCTCTGTCATTTCTTCAGCAGCTGTTGCCACACGACGCAATGTTACACAAGCT 74
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrLysLys 40
 |||||
 Db 75 ACATGGTTCACCTTCTACTGGTCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 134
 |||||

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 135 CAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 194
 |||||

QY 61 Trp 61
 |||||
 Db 195 TGG 197

RESULT 6
 TOMRBCSE
 LOCUS 742 bp mRNA linear PLN 27-APR-1993
 DEFINITION Tomato RuBP carboxylase small subunit LESS5 mRNA.
 ACCESSION M15236
 VERSION M15236.1 GI:170503
 KEYWORDS carboxylase; ribulose bisphosphate carboxylase.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
 AUTHORS McKnight, T.D., Alexander, D.C., Babcock, M.S. and Simpson, R.B.
 TITLE Nucleotide sequence and molecular evolution of two tomato genes
 encoding the small subunit of ribulose-1,5-bisphosphate carboxylase
 JOURNAL Gene 48 (1), 23-32 (1986)
 PUBMED 3557127
 COMMENT Original source text: Tomato (VF36), cDNA to mRNA, clone LESS5.
 Draft entry and computer-readable copy of sequence in [1] kindly
 provided by T.D.McKnight (15-APR-1987).

FEATURES
 source Location/Qualifiers
 1..742
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 <1..742
 /product="RuBP mRNA"
 28..570
 /note="ribulose-1,5-bisphosphate carboxylase, small
 subunit precursor"
 /codon_start=1
 /protein_id="AAA34192.1"
 /db_xref="GI:170504"
 /translation="MASSVITSSAAVATRSNVTQASVAPFTGLKSSATPVTKKQNLID
 ITSISANGRVSCMVQWPPINMKKYETLSVLPDLSEQLLSEVYLLKNGVPCLEFE
 TEGFVYRENNKSGFYDGRYTWKMLPMFGCTDTAQVLAEOEAKKAYPAQWVRIG
 FDNVRQVCISFIAYKPEGY"
 28..198
 /note="ribulose-1,5-bisphosphate carboxylase signal pept"
 199..567
 /product="ribulose-1,5-bisphosphate carboxylase"
 ORIGIN 614 bp upstream of Dnal site.
 Alignment Scores:
 Pred. No.: 5,27e-28 Length: 742
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSE (1-742)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||

Db 28 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 87
 QY 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrIysLys 40
 Db 89 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTTACAAGAAG 147
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db 148 CAAAACCTTGACATCACTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 207
 QY 61 TTP 61
 Db 208 TGG 210

RESULT 7
 BT013023
 LOCUS BT013023 1097 bp mRNA linear PLN 11-MAY-2004
 DEFINITION Lycopersicon esculentum clone 114270R, mRNA sequence.
 ACCESSION BT013023
 VERSION BT013023.1 GI:47104438
 KEYWORDS FLI CDNA.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 1097)
 AUTHORS Kirknes,E.F., Wang,W. and Vazeille,A.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
 Medical Center Drive, Rockville, MD 20850, USA
 FEATURES
 source
 1..1097
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /db_xref="taxon:4081"
 /clone="114270R"
 /tissue_type="Pseudomonas resistant"
 /note="TPRC058"

ORIGIN
 Alignment Scores:
 Pred. No.: 8,78e-28 Length: 1097
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x BT013023 (1-1097)

QY 1 MetaLaserSerValIleSerSerAlaLaValAlaThrArgSerAsnValThrGlnAla 20
 Db 362 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 421
 QY 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrIysLys 40
 Db 422 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTTACAAGAAG 481
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db 482 CAAAACCTTGACATCACTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 541
 QY 61 TTP 61
 Db 542 TGG 544

RESULT 8
 LERBCS2
 LOCUS LERBCS2 2776 bp DNA linear PLN 24-FEB-1999
 DEFINITION Tomato rbcS2 gene for ribulose 1,5-bisphosphate
 carboxylase/oxygenase small subunit.
 ACCESSION X05983

VERSION X05983.1 GI:19331
 KEYWORDS rbcS gene; ribulose bisphosphate carboxylase.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 2776)
 AUTHORS Sugita,M., Manzara,T., Pichersky,E., Cashmore,A. and Gruissem,W.
 TITLE Genomic organization, sequence analysis and expression of all five
 genes encoding the small subunit of ribulose-1,5-bisphosphate
 carboxylase/oxygenase from tomato
 Mol. Gen. Genet. 209 (2), 247-256 (1987)
 JOURNAL PUBMED 3478552
 REFERENCE 2 (bases 1 to 2776)
 AUTHORS Manzara,T.
 TITLE Direct Submission
 JOURNAL Submitted (02-FEB-1989)
 REFERENCE 3 (bases 1 to 2776)
 AUTHORS Manzara,T.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1989)
 COMMENT see X05982 - X05986.
 FEATURES
 Location/Qualifiers
 source
 1..2776
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="VENT Cherry LA1221"
 /db_xref="taxon:4081"
 /clone="Rbcs-2"
 enhancer 255..265
 CAAT_signal 274..278
 misc_feature 338..349
 TATA_signal 355..365
 exon 393..597
 gene 421..2495
 CDS join(421..597,1597..1731,1830..1882,2318..2495)
 /gene="rbcS2"
 /codon_start=1
 /product="ribulose 1,5-bisphosphate carboxylase/oxygenase"
 /protein_id="CAA29401.2"
 /db_xref="GI:4456641"
 /db_xref="GOA:P07179"
 /db_xref="UniProtKB/Swiss-Prot:P07179"
 /translation="MASSVSSAAVATRSNVTQASMVAPFTGLKSSATFPVTKQNLD
 ITSIASNGRVSQMVPPINMKKIETLSYLPDLSDBQLSEIEYLKNGWVPCLEPE
 TEHGFFVYRENNKSGYDGRYTWKLMKLPFGCTDATQVLAEVQAKKAYPQAWRIIG
 FDNVRQVCISFIAYKPEGY"
 421..591
 /gene="rbcS2"
 join(592..597,1597..1731,1830..1882,2318..2492)
 /gene="rbcS2"
 /product="ribulose 1,5-bisphosphate carboxylase/oxygenase"
 598..1596
 /number=1
 /gene="rbcS2"
 1597..1731
 /gene="rbcS2"
 1732..1829
 /gene="rbcS2"
 1739..1741
 /gene="rbcS2"
 /note="tca was taa in [2]"
 1830..1882
 /gene="rbcS2"
 1883..2317
 /number=3
 intron
 sig_peptide
 mat_peptide
 intron
 exon
 intron
 old_sequence
 exon
 intron


```

/ gene="rbcs2"
/ number=3
1956..1957
/ genes="rbcs2"
/ notes="aa was ada in [2]"
/ citation=[2]
2318..>2493
/ gene="rbcs2"
/ number=4
2637..2642

ORIGIN
Alignment Scores:
Pred. No.: 1,27e-26 Length: 2776
Score: 292.00 Matches: 61
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 1
Query Match: 95.7% Indels: 0
DB: Gaps: 0

US-10-718-311-15 (1-62) x LERBCS2 (1-2776)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 421 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTGCCACAGCAATGTTACACAAGCT 480

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 481 AGCATGGTTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 540

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 541 CAAAACCTTGACATCACTTCCATTTAGCAATGGTGGGAAGAGTTAGCTGCATGCAGGTA 600

QY 61 TrpHis 62
DB 601 ACTCAC 606

RESULT 9
AR528363
LOCUS AR528363 204 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6723895.
ACCESSION AR528363
VERSION AR528363.1 GI:53916426
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 204)
AUTHORS DeBonte,L.R. and Shorrosh,B.S.
TITLE Plants containing a cytosolic acetyl CoA-carboxylase nucleic acid
JOURNAL Patent: US 6723895-A 3 20-APR-2004;
Cargill, Incorporated; Wayzata, MN
FEATURES
source
1..204
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3,56e-25 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: Gaps: 0

US-10-718-311-15 (1-62) x AR528363 (1-204)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCACTTCTTCCTCTGCAGCAGTGGCCACCGCAGCAATGTTGCTCAAGCT 60

```

```

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AACATGGTTGCACCTTTCACCTGCGCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTTGCCAGCAACGGCGGAAGAGTGCATGTCAGGTG 180

QY 61 Trp 61
DB 181 TGG 183

RESULT 10
AR205973
LOCUS AR205973 499 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 41 from patent US 6369296.
ACCESSION AR205973
VERSION AR205973.1 GI:21503691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 499)
AUTHORS Ratcliff,F.Giles., Martin-Hernandez,A.Montserrat. and
Raulcombe,D.Charles.
TITLE Recombinant plant viral vectors
JOURNAL Patent: US 6369296-A 41 09-APR-2002;
FEATURES Location/Qualifiers
source
1..499
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,14e-24 Length: 499
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: Gaps: 0

US-10-718-311-15 (1-62) x AR205973 (1-499)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 8 ATGGCTTCCTCACTTCTTCCTCTGCAGCAGTGGCCACCGCAGCAATGTTGCTCAAGCT 67

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 68 AACATGGTTGCACCTTTCACCTGCGCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 127

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 128 CAAAACCTTGACATCACTTCCATTTGCCAGCAACGGCGGAAGAGTGCATGTCAGGTG 187

QY 61 Trp 61
DB 188 TGG 190

RESULT 11
NSRUB1
LOCUS NSRUB1 796 bp mRNA linear PLN 18-APR-2005
DEFINITION Nicotiana silvestris mRNA for the small subunit of ribulose
bisphosphate carboxylase.
ACCESSION X01722
VERSION X01722.1 GI:19759
KEYWORDS carboxylase; ribulose biphosphate carboxylase.
SOURCE Nicotiana sylvestris
ORGANISM Nicotiana sylvestris
REFERENCE 1 (bases 1 to 796)

```

AUTHORS Pinck, M., Guille, E., Durr, A., Hoff, M., Pinck, L. and Fleck, J.
 TITLE Complete sequence of one of the mRNAs coding for the small subunit
 of ribulose biphosphate carboxylase of Nicotiana sylvestris
 JOURNAL Biochimie 66 (7-8), 539-545 (1984)
 PUBMED 6549380
 COMMENT Data kindly reviewed (22-JUL-1985) by J. Fleck.

FEATURES

source
 1..796
 /organism="Nicotiana sylvestris"
 /mol_type="mRNA"
 /db_xref="taxon:4096"
 misc_feature
 1..60
 /note="5' untranslated region"
 repeat_region
 19..25
 /note="direct repeat 1"
 misc_feature
 21..46
 /note="pot. stem and loop structure"
 repeat_region
 27..33
 /note="direct repeat 1"
 repeat_region
 53..61
 /note="direct repeat 2"
 61..603
 /note="unnamed protein product; small subunit protein precursor"
 /codon_start=1
 /protein_id="CAA25862.1"
 /db_xref="GI:19760"
 /db_xref="GOA:P69250"
 /db_xref="InterPro:IPR000894"
 /db_xref="UniProtKB/Swiss-Prot:P69250"
 /translation="MASSVLSAAVATRSNVAQNMVAPFTGLKSAASFPVSRKQNL
 ITSASNGRQVQCMQVPPYINKKYETLSYLPDLSEQLLSEVEYLKNGWVPCLEFE
 TEGFVYRENNKSPGYDGRYWTWKLPMTGCTDATQVLAEEVEAKKAYPAQWIRIIG
 FDNVRQVCISFIAYKPEGY"
 transit_peptide
 61..231
 /note="transit peptide (aa -57 to -1)"
 mat_peptide
 232..600
 /product="mature small subunit protein (aa 1-123)"
 repeat_region
 249..257
 /note="direct repeat 2"
 misc_feature
 601..796
 /note="3' untranslated region"
 polyA_site
 796
 /note="polyA site"

ORIGIN

Alignment Scores:
 Pred. No.: 21e-24 Length: 796
 Score: 269.00 Matches: 53
 Percent Similarity: 96.7% Conservative: 6
 Best Local Similarity: 86.9% Mismatches: 2
 Query Match: 88.2% Indels: 0
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x NSRUB1 (1-796)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 Db 61 ATGGCTTCTCAGTTCTTCTTCCTGAGCAGTTGACCCGAGCAATGTTGCTCAAGCT 120
 Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 Db 121 AACATGGTTGACCTTTCACCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 180
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db 181 CAAACCTTGACATCACTTCCATTCAGCAACCGCGAAGAGTCAATGCATGCAGGTG 240

Qy 61 Trp 61
 Db 241 TGG 243

RESULT 12

AY220079
 LOCUS
 DEFINITION

Nicotiana tabacum putative ribulose biphosphate carboxylase small subunit protein precursor, mRNA, complete cds; nuclear gene for chloroplast product.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

AY220079.1 GI:30013662
 Nicotiana tabacum (common tobacco)
 Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 (bases 1 to 806)
 Tang, L., Dollins, C., Erra, S. and Petracek, M.
 Photosynthetic Control of Nuclear mRNA Gene Expression in Tobacco Unpublished

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

2 (bases 1 to 806)
 Tang, L., Dollins, C., Erra, S. and Petracek, M.
 Direct Submission
 Submitted (15-JAN-2003) Biochemistry and Molecular Biology, Oklahoma State University, 246 Noble Research Center, Stillwater, OK 74078, USA

FEATURES

source

1..806
 /organism="Nicotiana tabacum"
 /mol_type="mRNA"
 /db_xref="taxon:4097"
 61..803
 /note="similar to Nicotiana silvestris ribulose biphosphate carboxylase small subunit"
 /codon_start=1
 /product="putative ribulose biphosphate carboxylase small subunit protein precursor"
 /protein_id="AAP03874.1"
 /db_xref="GI:30013663"

CDS

/translation="MASSVLSAAVATRTNVAQNMVAPFTGLKSAASFPVSRKQNL
 ITSASNGRQVQCMQVPPYINKKYETLSYLPDLSEQLLSEIYLLKNGWVPCLEFE
 TERGFVYRENNKSPGYDGRYWTWKLPMTGCTDATQVLAEEVEAKKAYPAWIRIIG
 FDNVRQVCISFIAYKPEGY"

ORIGIN

Alignment Scores:
 Pred. No.: 5.14e-24 Length: 806
 Score: 266.00 Matches: 52
 Percent Similarity: 96.7% Conservative: 7
 Best Local Similarity: 85.2% Mismatches: 2
 Query Match: 87.2% Indels: 0
 DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x AY220079 (1-806)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 Db 61 ATGGCTTCTCAGTTCTTCTTCCTGAGCAGTTGCCACTGCACCAATGTTGCTCAAGCT 120
 Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 Db 121 AACATGGTTGACCTTTCACCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 180
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db 181 CAAACCTTGACATCACTTCCATTCAGCAATGTTGGAAGAGTCAATGCATGCAGGTA 240

Qy 61 Trp 61
 Db 241 TGG 243

RESULT 13

CVE18556
 LOCUS
 DEFINITION
 Cloning vector pSLJ8313, T-DNA region.
 ACCESSION Y18556

22574 bp DNA linear SYN 24-FEB-1999


```
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 644 AACATGGTTCACCTTCACTGCGCTTAAGTCAGCTCCTCATTCCTGTTTCAAGGAAG 703
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 704 CAAACCTTGACATCACTCCATTCGACGACGCGGAGAGAGTGCAATGCATGCAGGTA 763

RESULT 15
NPRBCS8B
LOCUS
DEFINITION Nicotiana plumbaginifolia rbcS-8B gene for ribulose biphosphate
ACCESSION X13711
VERSION X13711.1 GI:19705
KEYWORDS enhancer-like sequence; rbcS gene; ribulose biphosphate
carboxylase.
SOURCE Nicotiana plumbaginifolia (curled-leaved tobacco)
ORGANISM Nicotiana plumbaginifolia
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 2293)
REFERENCE Poulseu, C., Fluhr, R., Kauffman, J. M., Boutry, M. and Chua, N. H.
AUTHORS Characterization of an rbcS gene from Nicotiana plumbaginifolia and
TITLE expression of an rbcS-CAT chimeric gene in homologous and
heterologous nuclear background
Mol. Gen. Genet. 205, 193-200 (1986)
2 (bases 1 to 1130)
REFERENCE Poulseu, C. and Chua, N. H.
AUTHORS Dissection of 5' upstream sequences for selective expression of the
TITLE Nicotiana plumbaginifolia rbcS-8B gene
Mol. Gen. Genet. 214 (1), 16-23 (1988)
3226423
FEATURES
Location/Qualifiers
1..2293
/organism="Nicotiana plumbaginifolia"
/mol_type="genomic DNA"
/db_xref="taxon:4092"
/clone_lib="EMBL4"
1..450
/notes="enhancer-like sequence"
complement(403..410)
/notes="SV40 enhancer core similarity"
435..892
/notes="SV40 enhancer core similarity"
885..892
/notes="SV40 enhancer core similarity"
911..915
/notes="SV40 enhancer core similarity"
927..931
CAAT_signal
945..949
CAAT_signal
1008..1015
TATA_signal
1035
misc_feature
/notes="alternative transcriptional start"
1039..2171
precursor_RNA
1039..1289
/notes="primary transcript"
/number=1
1049..1051
misc_feature
/notes="alternative transcriptional start"
join(1113..1289,1381..1515,1644..1696,1815..1992)
CDS
/codon_start=1
/product="ribulose biphosphate carboxylase"
/protein_id="CAA31994.1"
/db_xref="GI:19706"
/db_xref="GOA:P26573"
/db_xref="UniProtKB/Swiss-Prot:P26573"
/translation="MASSVLSAAVATRSNVAQANVAPFTGLKSNASFPVSRKQNLID
ITSIASGRVQCMQWPPINKKYEITLSPDLSEQLLSIEYLLKSGWVPCLEFE
TERGFVYREHHSFGYDGRYWTMKLPMFGCTDQVLAEEVEAKAYPOAWVRRII
FDNRVQVCISFIAYKPEGY"
1290..1380
/number=1
intron
```

```
exon 1381..1515
intron /number=2
exon 1516..1643
intron /number=2
exon 1644..1696
intron /number=3
exon 1697..1814
intron /number=3
exon 1815..2171
misc_feature /number=4
misc_feature 2147..2151
polyA_signal /note="alternative transcriptional stop site"
misc_feature 2154..2159
misc_feature 2167..2170
misc_feature /note="alternative transcriptional stop site"
ORIGIN
Alignment Scores:
Pred. No.: 2 09e-22 Length: 2293
Score: 258.00 Matches: 52
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.7% Mismatches: 2
Query Match: 84.6% Indels: 0
DB: 4 Gaps: 0
US-10-718-311-15 (1-62) x NPRBCS8B (1-2293)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1113 ATGGCTTCCTCAGTTCTTCTCTGTCAGTCAGTGGCCAGCAGCAATGTTGCTCAAGCT 1172
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 1173 AACATGGTTCACCTTCACTGCGCTTAAGTCAGCTCCTCATTCCTGTTTCAAGGAAG 1232
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 1233 CAAACCTTGACATCACTCCATTCGACGACGCGGAGAGAGTGCAATGCATGCAGGTA 1292
```

Search completed: July 11, 2006, 21:35:35
Job time : 1360.71 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 20:58:40 ; Search time 36.8997 Seconds

(without alignments)
3276.788 Million cell updates/sec

Title: US-10-718-311-15

Perfect score: 305

Sequence: 1 MASSVISAAVATRNVTQA.....LDITSIASNGRVCQMWH 62

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 844069 seqs, 650066433 residues

Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/abs/ABSSWEB.spool/US10718311/runat_11072006.111535_5434/app.query.fasta.1
-DB=Published Applications NA New -OFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs06p
-USER=US10718311 @CGN 1.1 251 @runat_11072006.111535_5434 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

Published Applications NA New.*

1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PTCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269	88.2	762	8	US-11-247-587-125 Sequence 125, App
2	269	88.2	781	8	US-11-247-587-124 Sequence 124, App
3	250.5	82.1	698	8	US-11-247-587-128 Sequence 128, App
4	244.5	80.2	683	8	US-11-247-587-126 Sequence 126, App
5	242.5	79.5	700	8	US-11-247-587-127 Sequence 127, App
6	231	75.7	583	7	US-11-218-305-17631 Sequence 17631, A
7	227.5	74.6	727	8	US-11-247-587-129 Sequence 129, App
8	225	73.8	743	6	US-10-953-349-8894 Sequence 8894, App

9	224	73.4	658	7	US-11-218-305-17633 Sequence 17633, A
10	221	72.5	718	6	US-10-953-349-10521 Sequence 10521, A
11	215	70.5	640	7	US-11-218-305-17634 Sequence 17634, A
12	201	65.9	799	6	US-10-953-349-9569 Sequence 9569, App
13	201	65.9	828	7	US-11-218-305-17632 Sequence 17632, A
14	194	63.6	841	8	US-11-216-545-4173 Sequence 4173, App
15	194	63.6	1023	8	US-11-216-545-4174 Sequence 4174, App
16	183	60.0	1237	8	US-11-216-545-8381 Sequence 8381, App
17	131	43.0	947	6	US-10-449-902-24132 Sequence 24132, A
18	129	42.3	627	8	US-11-216-545-6801 Sequence 6801, App
19	129	42.3	882	8	US-11-266-748A-366894 Sequence 366894, App
20	129	42.3	882	8	US-11-266-748A-388679 Sequence 388679, App
21	129	42.3	882	8	US-11-266-748A-450273 Sequence 450273, App
22	125	41.0	845	6	US-10-449-902-19918 Sequence 19918, A
23	125	41.0	1999	6	US-10-449-902-17846 Sequence 17846, A
24	122	40.0	830	6	US-10-449-902-4174 Sequence 4174, App
25	120	39.3	899	6	US-10-449-902-18147 Sequence 18147, A
26	118	38.7	2409	6	US-10-449-902-20231 Sequence 20231, A
27	110	36.1	416	6	US-10-523-290-17 Sequence 17, Appl
28	110	36.1	5600	6	US-10-525-318-13 Sequence 13, Appl
29	110	36.1	7000	6	US-10-525-318-8 Sequence 8, Appl
30	72.5	23.8	3004	6	US-10-449-902-14773 Sequence 14773, A
31	69.5	22.8	964	8	US-11-266-748A-414508 Sequence 414508, App
32	66.5	21.8	1875	8	US-11-217-529-76380 Sequence 76380, A
33	66.5	21.8	4223	7	US-11-218-305-16072 Sequence 16072, A
34	65.5	21.5	1613	7	US-11-218-305-13081 Sequence 13081, A
35	64.5	21.1	1679	6	US-10-449-902-28323 Sequence 28323, A
36	64	21.0	1493	8	US-11-266-748A-185234 Sequence 185234, App
37	64	21.0	1493	8	US-11-266-748A-192750 Sequence 192750, App
38	63	20.7	1373	8	US-11-216-545-6828 Sequence 6828, App
39	63	20.7	1510	6	US-10-953-349-14286 Sequence 14286, A
40	63	20.7	2519	6	US-10-449-902-7755 Sequence 7755, App
41	63	20.7	3811	7	US-11-218-305-9560 Sequence 9560, App
42	63	20.7	3951	8	US-11-121-154-85 Sequence 85, Appl
43	62.5	20.5	1337	8	US-11-266-748A-368662 Sequence 368662, App
44	62.5	20.5	1337	8	US-11-266-748A-452041 Sequence 452041, App
45	62.5	20.5	1406	6	US-10-449-902-20240 Sequence 20240, A

ALIGNMENTS

RESULT 1

US-11-247-587-125

; Sequence 125, Application US/11247587

; Publication No. US20060130176A1

; GENERAL INFORMATION:

; APPLICANT: The Rockefeller University

; APPLICANT: Toboda, Jose Reyes

; APPLICANT: Zhang, Xiuren

; APPLICANT: Soyano, Takashi

; APPLICANT: Chua, Nam-Hai

; APPLICANT: Niu, Qi-Wen

; APPLICANT: Lin, Shih-Shun

; TITLE OF INVENTION: MicroRNAs

; FILE REFERENCE: 2312-124

; CURRENT APPLICATION NUMBER: US/11/247,587

; CURRENT FILING DATE: 2005-10-12

; PRIOR APPLICATION NUMBER: US 60/671,089

; PRIOR FILING DATE: 2005-04-14

; PRIOR APPLICATION NUMBER: PCT/US2004/033379

; PRIOR FILING DATE: 2004-10-12

; NUMBER OF SEQ ID NOS: 174

; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3

; SEQ ID NO 125

; LENGTH: 762

; TYPE: DNA

; ORGANISM: Nicotiana benthamiana

US-11-247-587-125

Alignment Scores:

Pred. No.:

Score: 2.86e-31 Length: 762

Matches: 269.00

Percent Similarity: 96.7% Conservative: 6

Best Local Similarity: 86.9% Mismatches: 2
 Query Match: 88.2% Indels: 0
 DB: 8 Gaps: 0

US-10-718-311-15 (1-62) x US-11-247-587-125 (1-762)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 Db ATGGCTTCTCAGTTCTTTCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 86
 Qy 21 SerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIlyss 40
 Db AACATGGTTGCACCTTCACTGGTCTTAAGTCAGCTGCTCGTCTTCAAGGAAG 146
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db CAAACACCTTGACATCACTTCCATTGCCAGCAACGCGGAAGAGTGCATGTCAGGTG 206
 Qy 61 Trp 61
 Db 207 TGG 209

RESULT 2

US-11-247-587-124
 ; Sequence 124, Application US/11247587
 ; Publication No. US20060130176A1

GENERAL INFORMATION:
 ; APPLICANT: The Rockefeller University
 ; APPLICANT: Toboda, Jose Reyes
 ; APPLICANT: Zhang, Xiuren
 ; APPLICANT: Soyano, Takashi
 ; APPLICANT: Chua, Nam-Hai
 ; APPLICANT: Niu, Qi-Wen
 ; APPLICANT: Lin, Shih-Shun
 ; TITLE OF INVENTION: MicroRNAs
 ; FILE REFERENCE: 2312-124
 ; CURRENT APPLICATION NUMBER: US/11/247,587
 ; PRIOR FILING DATE: 2005-10-12
 ; PRIOR APPLICATION NUMBER: US 60/671,089
 ; PRIOR FILING DATE: 2005-04-14
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379
 ; PRIOR FILING DATE: 2004-10-12
 ; NUMBER OF SEQ ID NOS: 174
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
 ; SEQ ID NO 124
 ; LENGTH: 781
 ; TYPE: DNA
 ; ORGANISM: Nicotiana benthamiana
 US-11-247-587-124

Alignment Scores:
 Pred. No.: 2,97e-31 Length: 781
 Score: 269.00 Matches: 53
 Percent Similarity: 96.7% Conservative: 6
 Best Local Similarity: 86.9% Mismatches: 2
 Query Match: 88.2% Indels: 0
 DB: 8 Gaps: 0

US-10-718-311-15 (1-62) x US-11-247-587-124 (1-781)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 Db ATGGCTTCTCAGTTCTTTCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 98
 Qy 21 SerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIlyss 40
 Db AACATGGTTGCACCTTTCACAGCTTTAAGTCAGCTGCTCGTCTTCAAGGAAG 158
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db CAAACACCTTGACATCACTTCCATTGCCAGCAACGCGGAAGAGTGCATGTCAGGTG 218
 Qy 61 Trp 61

Db 219 TGG 221

RESULT 3

US-11-247-587-128
 ; Sequence 128, Application US/11247587
 ; Publication No. US20060130176A1

GENERAL INFORMATION:
 ; APPLICANT: The Rockefeller University
 ; APPLICANT: Toboda, Jose Reyes
 ; APPLICANT: Zhang, Xiuren
 ; APPLICANT: Soyano, Takashi
 ; APPLICANT: Chua, Nam-Hai
 ; APPLICANT: Niu, Qi-Wen
 ; APPLICANT: Lin, Shih-Shun
 ; TITLE OF INVENTION: MicroRNAs
 ; FILE REFERENCE: 2312-124
 ; CURRENT APPLICATION NUMBER: US/11/247,587
 ; PRIOR FILING DATE: 2005-10-12
 ; PRIOR APPLICATION NUMBER: US 60/671,089
 ; PRIOR FILING DATE: 2005-04-14
 ; PRIOR APPLICATION NUMBER: PCT/US2004/033379
 ; PRIOR FILING DATE: 2004-10-12
 ; NUMBER OF SEQ ID NOS: 174
 ; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
 ; SEQ ID NO 128
 ; LENGTH: 698
 ; TYPE: DNA
 ; ORGANISM: Nicotiana benthamiana
 US-11-247-587-128

Alignment Scores:
 Pred. No.: 1.59e-28 Length: 698
 Score: 250.50 Matches: 52
 Percent Similarity: 91.9% Conservative: 5
 Best Local Similarity: 83.9% Mismatches: 4
 Query Match: 82.1% Indels: 1
 DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-128 (1-698)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGln 19
 Db ATGGCTTCTCAGTTCTTTCTCAGCTGCTGCCACCGCCTTAATGCTGCTCAA 79
 Qy 20 AlaSerMetValAlaProPheThrGlyLeuIlyssSerSerAlaThrPheProValThrIly 39
 Db GCCAGCATGGTTGCACCTTCACTGGCTCAATCTGCTCTCTCTCTTACCAGA 139
 Qy 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
 Db AACAAAAACCTTGACATTACATTCATTGCTAGCAATGTTGGAAGAGTCCAATGTCATG 199
 Qy 60 ValTrp 61
 Db 200 GTGTGG 205

RESULT 4

US-11-247-587-126
 ; Sequence 126, Application US/11247587
 ; Publication No. US20060130176A1

GENERAL INFORMATION:
 ; APPLICANT: The Rockefeller University
 ; APPLICANT: Toboda, Jose Reyes
 ; APPLICANT: Zhang, Xiuren
 ; APPLICANT: Soyano, Takashi
 ; APPLICANT: Chua, Nam-Hai
 ; APPLICANT: Niu, Qi-Wen
 ; APPLICANT: Lin, Shih-Shun
 ; TITLE OF INVENTION: MicroRNAs
 ; FILE REFERENCE: 2312-124
 ; CURRENT APPLICATION NUMBER: US/11/247,587
 ; PRIOR FILING DATE: 2005-10-12


```
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 126
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-126

Alignment Scores:
Pred. No.: 1,24e-27 Length: 683
Score: 244.50 Matches: 50
Percent Similarity: 91.9% Conservatives: 7
Best Local Similarity: 80.6% Mismatches: 4
Query Match: 80.2% Indels: 1
DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-126 (1-683)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
Db 37 ATGGCTTCTCAGTTATGCTCAGCAGCTGCTGTGGACCGGCCAATGCTGCTCAA 96
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39
Db 97 GCCAACATGGTGCACCTTCACTGGCCTCAAGTCGGCTCCTCCTCTCCCTGTTACCAGG 156
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 157 AAACAAAACCTTGACATTACCTCCATTGCTAGCAATGGTGAAGAGTTCAATGCATGCAG 216
QY 60 ValTtp 61
Db 217 GTGTGG 222

RESULT 5
US-11-247-587-127
; Sequence 127, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; PRIOR FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-127

Alignment Scores:
Pred. No.: 2.58e-27 Length: 700
Score: 242.50 Matches: 50
Percent Similarity: 90.3% Conservatives: 6
Best Local Similarity: 80.6% Mismatches: 5
Query Match: 79.5% Indels: 1
DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-127 (1-700)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
Db 2 ATGGCTTCTCAGTTATGCTCAGCAGCTGCTGTGGACCGGCCAATGCTGCTCAA 61
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39
Db 62 GCCAGTATGGTTGCACCTTTCACCTGGCCTCAAGTCGGCAACCTCCTTCCCTGTTCCAGA 121
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 122 AAACAAAACCTTGACATTACCTCCATTGCTAGCAACCGCGGAGAGTTCAATGCATGCAG 181
QY 60 ValTtp 61
Db 182 GTGTGG 187

RESULT 6
US-11-218-305-17631
; Sequence 17631, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Mc Laird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17631
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17631

Alignment Scores:
Pred. No.: 1.09e-25 Length: 583
Score: 231.00 Matches: 48
Percent Similarity: 88.5% Conservatives: 6
Best Local Similarity: 78.7% Mismatches: 5
Query Match: 75.7% Indels: 2
DB: 7 Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17631 (1-583)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 33 ATGGCTTCTCAGTTATGCTCCTCCCGCGCTGGGTACC-----TCACCGGCTCAAGCC 86
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 87 ACCATGGTTGCTCCATTACCGCGCTTGAAGTCATCCGCTGCATTCCTCCAGTCACCCGCAAA 146
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 147 AGCAACACCTGATATTACTTCCATTGCAAGCAACGGAGGAAGAGTTAGTGCATGAAGGTG 206
QY 61 Ttp 61
Db 207 TGG 209

RESULT 7
US-11-247-587-129
; Sequence 129, Application US/11247587
; Publication No. US20060130176A1
```

```

; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 129
; LENGTH: 727
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-129

Alignment Scores:
Pred. No.: 5,068-25 Length: 727
Score: 227.50 Matches: 46
Percent Similarity: 90.2% Conservatives: 9
Best Local Similarity: 75.4% Mismatches: 5
Query Match: 74.6% Indels: 1
DB: 8 Gaps: 1

US-10-718-311-15 (1-62) x US-11-247-587-129 (1-727)

Qy 2 AlaSerValIleSerSer---AlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 GCTTCTCAGTTATGCTCTCAGCTCGCGTGTTCACCGCGCGCAATGCTGTTCAGGCC 67
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 68 ACATGGTGGCGACCTTCACTGGCTCAAGCGCGCTCTCTCCCGGTTCCAGGAAA 127
Qy 41 GlnAsnLeuAspPileThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAACCTTGACATTACTTCCATTCTAGAAATGTTGGAAGAGTCCAATGCATGCAGCGG 187
Qy 61 Trp 61
Db 188 TGG 190

RESULT 8
; Sequence 8894, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8894
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8894

Alignment Scores:
Pred. No.: 1,258-24 Length: 743
Score: 225.00 Matches: 47
Percent Similarity: 86.9% Conservatives: 6

; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaIRD, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17633
; LENGTH: 658
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17633

Alignment Scores:
Pred. No.: 1,498-24 Length: 658
Score: 224.00 Matches: 47
Percent Similarity: 86.9% Conservatives: 6
Best Local Similarity: 77.0% Mismatches: 6
Query Match: 73.4% Indels: 2
DB: 7 Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17633 (1-658)

Qy 1 MetAlaSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 45 ATGGCTTCTCTATGCTTCTCTCGCGTGTGTTTAC-----TCCCGCGCTCAAGCC 98
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 99 ACCATGGTGGCTCCATTCACCGGCTTGAAGTCATCCGCTGCTATTCACCGCGCAA 158
Qy 41 GlnAsnLeuAspPileThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 159 GCAACCAACGACATTACTTCCATTGCAAGCAACGGAAGAGTTAGTGTGATGAAGGTG 218
Qy 61 Trp 61
Db 219 TGG 221

RESULT 10

```

```
US-10-953-349-10521
; Sequence 10521, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; PRIOR FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10521
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10521

Alignment Scores:
Pred. No.: 4,79e-24 Length: 718
Score: 221.00 Matches: 46
Percent Similarity: 86.9% Conservative: 7
Best Local Similarity: 75.4% Mismatches: 6
Query Match: 72.5% Indels: 2
DB: 6 Gaps: 1

US-10-718-311-15 (1-62) x US-10-953-349-10521 (1-718)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 32 ATGGCTTCTCTATGCTCTCTCCGCGCTGTGTTACC-----TCCCGGGCTCAAGCC 85
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 86 ACCATGGTCGCTCCATTCACTGGTTTGAAGTCATCGCTTCTTCCCGGTCAACCCGCAAG 145
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 146 GCCAACACGACATTACTTCCATCAACGAATGGGGGAAGAGTTAGTCGCATGAAGGTG 205
QY 61 TTP 61
Db 206 TGG 208

RESULT 11
US-11-218-305-17634
; Sequence 17634, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17634
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-17634

Alignment Scores:
Pred. No.: 3,28e-23 Length: 640
Score: 215.00 Matches: 45
Percent Similarity: 83.6% Conservative: 6
Best Local Similarity: 73.8% Mismatches: 8
Query Match: 70.5% Indels: 2
DB: 1 Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17634 (1-640)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 33 ATGGCTTACTCTATGCTCTCTCCGCGCTGTGTTACC-----TCCCGGGCTCAAGCC 86
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 87 ACCATGGTCGCTCCATTCAACCGCTTGAAGTCTTCTCTGCATTCCAGTCACCCGCAAG 146
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 147 GCCAACACGACATTACTTCCATCGTAGCAACGGAGGAAGAGTTAGTCGCATGCAAGGTG 206
QY 61 TTP 61
Db 207 TGG 209

RESULT 12
US-10-953-349-9569
; Sequence 9569, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9569
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9569

Alignment Scores:
Pred. No.: 5,9e-21 Length: 799
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: 6 Gaps: 1

US-10-718-311-15 (1-62) x US-10-953-349-9569 (1-799)
QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 47 ATGGCTTCTCTATGCTCTCTCCGCTACTAGTTGTT-----GCCCTCTCCGGCTCAGGCC 100
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 101 ACTATGGTCGCTCCITTCACGAGCTTAAGTCTCTCGCTGCCTTCCAGCCACCCGCAAG 160
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 161 GCTAACACGACATTACTTCCATCAACGAATGGGGGAAGAGTTAACTGCATGCAAGGTG 220
QY 61 TTP 61
Db 221 TGG 223

RESULT 13
US-11-218-305-17632
; Sequence 17632, Application US/11218305
```

; Publication No. US20060141495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MONSANTO TECHNOLOGY, LLC
 ; APPLICANT: McLaird, Paul L.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Wu, Kunsheng
 ; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
 ; FILE REFERENCE: 38-21 (53660)B
 ; CURRENT APPLICATION NUMBER: US/11/218,305
 ; CURRENT FILING DATE: 2005-09-01
 ; PRIOR APPLICATION NUMBER: US 60/606,880
 ; PRIOR FILING DATE: 2004-09-01
 ; NUMBER OF SEQ ID NOS: 25043
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 17632
 ; LENGTH: 828
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 US-11-218-305-17632

Alignment Scores:
 Pred. No.: 6,218-21 Length: 828
 Score: 201.00 Matches: 41
 Percent Similarity: 80.3% Conservative: 8
 Best Local Similarity: 67.2% Mismatches: 10
 Query Match: 65.9% Indels: 2
 DB: 7 Gaps: 1

US-10-718-311-15 (1-62) x US-11-218-305-17632 (1-828)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||.....
 Db 50 ATGGCTTCTCTATGCTCTTCGCTACTATGGTT-----GCTCTCGGCTCAGGCC 103
 |||||.....
 Qy 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrLysLys 40
 ::::|.....
 Db 104 ACTATGGTGGCTCTTCAACGGACTTAAGTCTCGCTGCCCTCCAGCCACCGCAAG 163
 |||||.....
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||.....
 Db 164 GCTAACACGACATTACTTCCATCAACGCGGGAAGAGTTAACTGATGCAGGTG 223
 |||||.....

Qy 61 Trp 61

Db 224 TGG 226

RESULT 14

; US-11-216-545-4173
 ; Sequence 4173, Application US/11216545
 ; Publication No. US20060135758A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MONSANTO Technology, LLC
 ; APPLICANT: McLaird, Paul L.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Wu, Kunsheng
 ; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
 ; FILE REFERENCE: 38-21 (53659)B
 ; CURRENT APPLICATION NUMBER: US/11/216,545
 ; CURRENT FILING DATE: 2005-08-31
 ; PRIOR APPLICATION NUMBER: US 60/606,062
 ; PRIOR FILING DATE: 2004-08-31
 ; NUMBER OF SEQ ID NOS: 8783
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4173
 ; LENGTH: 841
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-11-216-545-4173

Alignment Scores:
 Pred. No.: 7,278-20 Length: 841

Score: 194.00 Matches: 45
 Percent Similarity: 77.0% Conservative: 2
 Best Local Similarity: 73.8% Mismatches: 12
 Query Match: 63.6% Indels: 2
 DB: 8 Gaps: 2

US-10-718-311-15 (1-62) x US-11-216-545-4173 (1-841)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||.....
 Db 37 ATGGCTTCTCTCAATGATCTCTCCACAGCTGTACCACC---GTCAACCGTCCCGGTGCC 93
 |||||.....
 Qy 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrLysLys 40
 ::::|.....
 Db 94 GGCATGGTTGCTCCATTACCGGGCTCAATCCATGGCTGGCTTCCCC---ACGAGGAAG 150
 |||||.....
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||.....
 Db 151 ACCAACCAATGACATTACCTCCATTGCTAGCAACGCTGGAAGAGTACAATGCAATGCAGGTG 210
 |||||.....
 Qy 61 Trp 61
 |||||.....
 Db 211 TGG 213

RESULT 15

; US-11-216-545-4174
 ; Sequence 4174, Application US/11216545
 ; Publication No. US20060135758A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MONSANTO Technology, LLC
 ; APPLICANT: McLaird, Paul L.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Wu, Kunsheng
 ; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
 ; FILE REFERENCE: 38-21 (53659)B
 ; CURRENT APPLICATION NUMBER: US/11/216,545
 ; CURRENT FILING DATE: 2005-08-31
 ; PRIOR APPLICATION NUMBER: US 60/606,062
 ; PRIOR FILING DATE: 2004-08-31
 ; NUMBER OF SEQ ID NOS: 8783
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 4174
 ; LENGTH: 1023
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; NAME/KEY: misc_feature
 ; LOCATION: (12)..(12)
 ; OTHER INFORMATION: n is a, c, g, or t
 US-11-216-545-4174

Alignment Scores:
 Pred. No.: 9,628-20 Length: 1023
 Score: 194.00 Matches: 45
 Percent Similarity: 77.0% Conservative: 2
 Best Local Similarity: 73.8% Mismatches: 12
 Query Match: 63.6% Indels: 2
 DB: 8 Gaps: 2

US-10-718-311-15 (1-62) x US-11-216-545-4174 (1-1023)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||.....
 Db 228 ATGGCTTCTCTCAATGATCTCTCCACAGCTGTACCACCT---GTCAACCGTCCCGGTGCC 284
 |||||.....
 Qy 21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrLysLys 40
 ::::|.....
 Db 285 GGCATGGTTGCTCCATTACCTCAAGTCCATGGCTGGCTTCCCC---ACGAGGAAG 341
 |||||.....
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||.....
 Db 342 ACCAACCAATGACATTACCTCCATTGCTAGCAACGCTGGAAGAGTGAATGCAATGCAGGTG 401
 |||||.....

Qy 61 Trip 61
Db 402 TGG 404

Search completed: July 11, 2006, 21:05:59
Job time : 37.8997 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: July 11, 2006, 20:23:26 ; Search time 452.45 Seconds
(without alignments)
2525.695 Million cell updates/sec
Title: US-10-718-311-15
Perfect score: 305
Sequence: 1 MASSVISAANVATRSNVTA.....LDITSIASNGRVSCHQVWH 62
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 18892170 seqs, 6143817638 residues
Total number of hits satisfying chosen parameters: 37784340
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Command line parameters:
-Q=/abes/ABSSWEB spool/US10718311/runat 11072006 111531 5371/app_query.fasta_1
-DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss08
-USER=US10718311 @CN 1 1 2326 @runat 11072006 111531 5371 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:
1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
5: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
9: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
10: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
11: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
12: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
13: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
14: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
15: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
16: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	305	100.0	684	3	US-09-855-341-7 Sequence 7, Appli

2	305	100.0	684	3	US-09-896-8668-16
3	305	100.0	684	7	US-10-359-369-41
4	305	100.0	684	8	US-10-699-050-16
5	305	100.0	684	8	US-10-718-311-7
6	305	100.0	684	9	US-10-462-162-33
7	269	88.2	204	3	US-09-839-477-3
8	269	88.2	204	8	US-10-758-064-3
9	269	88.2	297	10	US-10-487-901-7209
10	269	88.2	377	10	US-10-487-901-3504
11	269	88.2	489	10	US-10-487-901-3503
12	269	88.2	504	10	US-10-487-901-3207
13	269	88.2	718	10	US-10-487-901-5765
14	269	88.2	736	10	US-10-487-901-1704
15	269	88.2	736	10	US-10-487-901-5043
16	265	86.9	668	10	US-10-487-901-7208
17	253.5	83.1	704	9	US-10-425-115-85509
18	250.5	82.1	683	10	US-10-487-901-1713
19	250.5	82.1	737	10	US-10-487-901-5045
20	250.5	82.1	737	10	US-10-487-901-5766
21	244.5	80.2	738	10	US-10-487-901-1716
22	244.5	80.2	753	10	US-10-487-901-5764
23	244.5	80.2	754	10	US-10-487-901-5028
24	244	80.0	174	3	US-09-854-286-15
25	242.5	79.5	714	10	US-10-487-901-1703
26	242.5	79.5	736	10	US-10-487-901-5044
27	235	77.0	169	6	US-10-165-420-3
28	232.5	76.2	654	10	US-10-487-901-5042
29	225	73.8	385	3	US-09-732-627A-1176
30	225	73.8	724	8	US-09-910-664-55
31	225	73.8	724	8	US-10-333-184-54
32	225	73.8	726	8	US-10-333-184-202
33	225	73.8	771	3	US-09-770-445-918
34	225	73.8	1205	3	US-09-770-445-41
35	221	72.5	363	3	US-09-770-791-468
36	221	72.5	711	3	US-09-910-664-32
37	221	72.5	711	8	US-10-333-184-31
38	221	72.5	739	8	US-10-333-184-182
39	221	72.5	1036	9	US-10-739-930-2595
40	217	71.1	767	3	US-09-770-445-931
41	217	71.1	780	3	US-09-770-445-880
42	214	70.2	362	3	US-09-770-791-463
43	214	70.2	392	3	US-09-770-423-971
44	213.5	70.0	424	10	US-10-487-901-1710
45	213.5	70.0	424	10	US-10-487-901-5032

ALIGNMENTS

RESULT 1
US-09-855-341-7
; Sequence 7, Application US/09855341
; Patent No. US2002002715A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DEW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-09-855-341-7

Alignment Scores: 8.14e-36 Length: 684
Pred. No.: 684


```
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protein
US-10-699-050-16

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-718-311-15 (1-62) x US-10-699-050-16 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCCTGCTCAATCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTCCATTGCTAGCAATGGTGGAGAGTTAGTTCATGCATGCAGGTG 180

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-718-311-15 (1-62) x US-10-699-050-16 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCCTGCTCAATCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTCCATTGCTAGCAATGGTGGAGAGTTAGTTCATGCATGCAGGTG 180

QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 5
US-10-718-311-7
; Sequence 7, Application US/10718311
; Publication No. US20040143867A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUIT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/10/718,311
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/855,341
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-10-718-311-7

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             8          Gaps:      0

US-10-718-311-15 (1-62) x US-10-718-311-7 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCCTGCTCAATCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTCCATTGCTAGCAATGGTGGAGAGTTAGTTCATGCATGCAGGTG 180

QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 6
US-10-462-162-33
; Sequence 33, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462,162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-462-162-33

Alignment Scores:
Pred. No.:      8,14e-36      Length:      684
Score:          305.00      Matches:      62
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:    100.0%      Indels:      0
DB:             9          Gaps:      0

US-10-718-311-15 (1-62) x US-10-462-162-33 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCCTGCTCAATCTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTCCATTGCTAGCAATGGTGGAGAGTTAGTTCATGCATGCAGGTG 180

QY 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 7
US-09-839-477-3
; Sequence 3, Application US/09839477
; Publication No. US20030167523A1
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
```

```
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
US-09-839-477-3

Alignment Scores:
Pred. No.: 4,62e-31 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-15 (1-62) x US-09-839-477-3 (1-204)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTTCTTCTCTCGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCTCTGTTTCAAGGAAG 120
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGTCAGGTG 180
Qy 61 Trp 61
Db 181 TGG 183

RESULT 8
US-10-758-064-3
; Sequence 3, Application US/10758064
; Publication No. US20040128718A1
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/10/758,064
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US/09/839,477
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
US-10-758-064-3

Alignment Scores:
Pred. No.: 4,62e-31 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 8 Gaps: 0
```

```
US-10-718-311-15 (1-62) x US-10-758-064-3 (1-204)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTTCTTCTCTCGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCTCTGTTTCAAGGAAG 120
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGTCAGGTG 180
Qy 61 Trp 61
Db 181 TGG 183

RESULT 9
US-10-487-901-7209
; Sequence 7209, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McGreery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7209
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-7209

Alignment Scores:
Pred. No.: 7,79e-31 Length: 297
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-7209 (1-297)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 97 ATGGCTTCTCAGTTCTTCTCTCGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 156
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 157 AACATGGTTGCACCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCTCTGTTTCAAGGAAG 216
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyArgValSerCysMetGlnVal 60
Db 217 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGTCAGGTG 276
Qy 61 Trp 61
Db 181 TGG 183
```

```

; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3503
; LENGTH: 489
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
;
US-10-487-901-3503

Alignment Scores:
Pred. No.: 1.56e-30 Length: 489
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-3503 (1-489)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
   |||||
Db 79 ATGGCTTCCTCAGTCTTCCTCAGCAGCAGTTCGCCACCGCAGCAATGTTGCTCAAGCT 138

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
   ::|||
Db 139 AACATGGTTGCACCTTTTCACAGGTCTTAAAGTGTGCTGCTCCTCANTCCCTGTTTCAAGAAAG 198

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
   |||||
Db 199 CAAACCTTGACATCACTTCATTCATGCCAGCAACGGCGGAGAGTGCATATGCATGCAGGTG 258

Qy 61 Trp 61
   |||
Db 259 TGG 261

RESULT 12
US-10-487-901-7207
; Sequence 7207, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Peill, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7207
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
;
US-10-487-901-7207

Alignment Scores:
Pred. No.: 1.62e-30 Length: 504

```

Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-7207 (1-504)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 78 ATGGCTTCTCAGTTCCTTCTCAGCAGCAGTTCACCCGAGCAATGTTGCTCAAGCT 137
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 138 AACATGGTTGCACCTTTACAGGCTTTAAGTCTGTGCTCATTCCTGTTTCAAGAAG 197
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 198 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 257
Qy 61 Trp 61
Db 258 TGG 260

RESULT 13

US-10-487-901-5765
; Sequence 5765, Application US/10487901
; Publication No. US20050091708A1

GENERAL INFORMATION:

; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOM-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5765

LENGTH: 718

TYPE: DNA

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: Synthetic

US-10-487-901-5765

Alignment Scores:

Pred. No.: 2,666-30 Length: 718
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-5765 (1-718)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 45 ATGGCTTCTCAGTTCCTTCTCAGCAGCAGTTCACCCGAGCAATGTTGCTCAAGCT 104
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 105 AACATGGTTGCACCTTTACAGGCTTTAAGTCAAGTGCCTGCTGCTCCCTGTTTCAAGAAG 164
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60

Db 165 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 224
Qy 61 Trp 61
Db 225 TGG 227

RESULT 14

US-10-487-901-1704
; Sequence 1704, Application US/10487901
; Publication No. US20050091708A1

GENERAL INFORMATION:

; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOM-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1704
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic

US-10-487-901-1704

Alignment Scores:

Pred. No.: 2,756-30 Length: 736
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-1704 (1-736)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 63 ATGGCTTCTCAGTTCCTTCTCAGCAGCAGTTCACCCGAGCAATGTTGCTCAAGCT 122
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 123 AACATGGTTGCACCTTTACAGGCTTTAAGTCAAGTGCCTGCTGCTCCCTGTTTCAAGAAG 182
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 183 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 242
Qy 61 Trp 61
Db 243 TGG 245

RESULT 15

US-10-487-901-5043
; Sequence 5043, Application US/10487901
; Publication No. US20050091708A1

GENERAL INFORMATION:

; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus

; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Croxley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5043
; LENGTH: 736
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-5043

Alignment Scores:
Pred. No.: 2.75e-30 Length: 736
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x US-10-487-901-5043 (1-736)

Qy	1	MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla	20
Db	63	ATGGCTTCCTCAGTTCTTTCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT	122
Qy	21	SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys	40
Db	123	AACATGGTTGCACCTTCTACTGGTCTTAAGTCAGCTGCCCTGCTTCCTGTTTCAAGGAAG	182
Qy	41	GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal	60
Db	183	CAAAACCTTGACATCACTTCCATTGCCAGCAACGCGGAAGAGTGCAATGCATGCAGGTG	242
Qy	61	Trp	61
Db	243	TGG	245

Search completed: July 11, 2006, 22:12:04
Job time : 452.45 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 20:16:09 ; Search time 54.7059 Seconds
(without alignments)
3180.885 Million cell updates/sec

Title: US-10-718-311-15
Perfect score: 305
Sequence: 1 MASSVISNAVATRSNTQA.....LDITSIASNGRVSCMQVWH 62

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10718311/runat 11072006 111527 5281/app query.fasta_1
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs06p -USER=US10718311@CN 1 1 381 @runat 11072006 111527 5281
-NCPUS=3 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:**

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq:
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq:
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	305	100.0	684	3	US-09-896-866B-16
2	305	100.0	684	3	US-09-855-341-7
3	289	88.2	204	3	US-09-839-477-3
4	269	88.2	499	3	US-09-495-797-41
5	254	83.3	177	2	US-08-095-726-12
6	254	83.3	177	2	US-08-096-043-12
7	254	83.3	177	2	US-08-093-577-8
8	254	83.3	177	2	US-08-331-004A-6

9	254	83.3	177	2	US-08-096-623A-12	Sequence 12, Appl
10	254	83.3	177	7	PCT-US95-13937A-6	Sequence 6, Appl
11	249	81.6	1442	2	US-08-152-483B-8	Sequence 8, Appl
12	201	65.9	264	3	US-09-441-340-9	Sequence 9, Appl
13	201	65.9	268	3	US-09-186-002-5	Sequence 5, Appl
14	201	65.9	279	2	US-08-391-339-9	Sequence 9, Appl
15	201	65.9	279	2	US-08-484-274A-9	Sequence 9, Appl
16	201	65.9	279	9	US-09-612-404-9	Sequence 9, Appl
17	201	65.9	355	2	US-08-090-523-5	Sequence 5, Appl
18	201	65.9	355	2	US-08-398-627-5	Sequence 5, Appl
19	201	65.9	355	2	US-08-406-858-5	Sequence 5, Appl
20	201	65.9	355	2	US-08-476-519-7	Sequence 7, Appl
21	201	65.9	355	2	US-08-120-703A-5	Sequence 5, Appl
22	201	65.9	355	3	US-08-399-023-5	Sequence 5, Appl
23	201	65.9	355	7	PCT-US91-04036-5	Sequence 5, Appl
24	201	65.9	355	7	PCT-US94-05275-5	Sequence 5, Appl
25	201	65.9	355	7	PCT-US95-09323-7	Sequence 7, Appl
26	201	65.9	509	3	US-09-011-151-1	Sequence 1, Appl
27	201	65.9	696	3	US-09-441-340-11	Sequence 11, Appl
28	201	65.9	1178	3	US-09-434-039A-36	Sequence 36, Appl
29	201	65.9	1224	3	US-09-434-039A-34	Sequence 34, Appl
30	201	65.9	1630	3	US-09-441-340-23	Sequence 23, Appl
31	201	65.9	8012	3	US-09-182-117-1	Sequence 1, Appl
32	201	65.9	8012	3	US-09-434-039A-1	Sequence 1, Appl
33	201	65.9	8418	3	US-09-182-117-5	Sequence 5, Appl
34	201	65.9	8418	3	US-09-434-039A-5	Sequence 5, Appl
35	201	65.9	8798	3	US-09-182-117-4	Sequence 4, Appl
36	201	65.9	8798	3	US-09-434-039A-4	Sequence 4, Appl
37	201	65.9	10339	3	US-09-186-002-13	Sequence 13, Appl
38	201	65.9	10846	3	US-09-098-219B-5	Sequence 5, Appl
39	201	65.9	10846	3	US-10-164-204-5	Sequence 5, Appl
40	201	65.9	10846	3	US-09-923-109-5	Sequence 5, Appl
41	194	63.6	281	2	US-08-469-202-15	Sequence 15, Appl
42	194	63.6	281	2	US-08-469-202-17	Sequence 17, Appl
43	194	63.6	281	2	US-08-469-202-19	Sequence 19, Appl
44	194	63.6	281	2	US-08-484-434C-15	Sequence 15, Appl
45	194	63.6	281	2	US-08-484-434C-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-09-896-866B-16

; Sequence 16, Application US/09896866B

; Patent No. 6642036

; GENERAL INFORMATION:

; APPLICANT: Flint, Dennis

; APPLICANT: Meyer, Knut

; APPLICANT: Viitanen, Paul

; TITLE OF INVENTION: Sinapoylgucose:Malate Sinapoyltransferase Form Malate Conjugate

; FILE REFERENCE: BC1034 US NA

; CURRENT APPLICATION NUMBER: US/09/896,866B

; CURRENT FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/216,615

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16

; LENGTH: 684

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei

US-09-896-866B-16

Alignment Scores:	7,84e-35	Length:	684
Pred. No.:	305.00	Matches:	62
Score:	100.0%	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	100.0%		

```
DB: 3 Gaps: 0
US-10-718-311-15 (1-62) x US-09-896-866B-16 (1-684)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCTATTCTTCAGCAGCTGTGGCCACAGCGAGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGCACTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGTG 180
Qy 61 TrpHis 62
Db 181 TGGCAT 186
RESULT 2
US-09-855-341-7
; Sequence 7, Application US/09855341
; Patent No. 6683231
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; FILE REFERENCE: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic CPL
US-09-855-341-7
Alignment Scores:
Pred. No.: 7,84e-35 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-15 (1-62) x US-09-855-341-7 (1-684)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCTATTCTTCAGCAGCTGTGGCCACAGCGAGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGCACTTCACTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAAGTG 180
Qy 61 TrpHis 62
Db 181 TGGCAT 186
RESULT 3
US-09-839-477-3
; Sequence 3, Application US/09839477
; Patent No. 6723895
```

```
; GENERAL INFORMATION:
; APPLICANT: Shorirosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: COA-CARBOXYLASE
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
US-09-839-477-3
Alignment Scores:
Pred. No.: 2,32e-30 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-15 (1-62) x US-09-839-477-3 (1-204)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCTATTCTTCAGCAGCTGTGGCCACAGCGAGCAATGTTCTCAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCACTTTCACCTTCAAGCTTAAGTCAGCTGCTCCTCCTCCCTGTTCAAGGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCACTTCCATTGCCAGCAACGCGGAGAGTGCATGCAATGCAAGGTG 180
Qy 61 Trp 61
Db 181 TGG 183
RESULT 4
US-09-495-797-41
; Sequence 41, Application US/09495797
; Patent No. 6369296
; GENERAL INFORMATION:
; APPLICANT: Ratcliff, Frank G
; APPLICANT: Martin-Hernandez, Ana M
; APPLICANT: Baulcombe, David C
; TITLE OF INVENTION: Viral Vectors
; FILE REFERENCE: Mewburn 43,047
; CURRENT APPLICATION NUMBER: US/09/495,797
; CURRENT FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N. benthamiana rubisco partial cDNA sequence
US-09-495-797-41
Alignment Scores:
Pred. No.: 8,49e-30 Length: 499
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
```


Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-15 (1-62) x US-09-495-797-41 (1-499)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAenValThrGlnAla 20
Db 8 ATGGCTTCTCAGTTCTTCTCCCTGCGAGAGTTGCCACCGCAGCAATGTTGCTCAAGCT 67
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 68 AACATGGTTGCACCTTCTACCTGCGCTTAAGTCAGTCGCTCATTCCTCTGTTTCAAGGAAG 127
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAACCTTGACATCACTTCCATTCGTCAGCAACCGCGGAAGAGTGAATGCATGCAGGTG 187
QY 61 TTP 61
Db 188 TGG 190

RESULT 5

US-08-095-726-12
; Sequence 12, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huel-Che B
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-095-726-12

Alignment Scores:
Pred. No.: 2,87e-28 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6

Best Local Similarity: 86.4% Mismatches: 2
Query Match: 83.3% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x US-08-095-726-12 (1-177)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAenValThrGlnAla 20
Db 1 ATGGCTTCTCAGTTCTTCTCCCTGCGAGAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTCTACCTGCGCTTAAGTCAGTCGCTCATTCCTCTGTTTCAAGGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCACTTCCATTCGTCAGCAACCGCGGAAGAGTGAATGCATGCAG 177

RESULT 6

US-08-096-043-12
; Sequence 12, Application US/08096043
; Patent No. 5530189
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huel-Che B
; TITLE OF INVENTION: Lycopene Biosynthesis in
; Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,043
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,568
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530189val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

US-08-096-043-12

Alignment Scores:
Pred. No.: 2,87e-28 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 83.3% Indels: 0
DB: 2 Gaps: 0

Db 121 CAAAACCTTGACATCACTCCATTCGACGACGCGGAGAGTGCATGATGCAG 177

RESULT 9

US-08-096-623A-12

; Sequence 12, Application US/08096623A

; Patent No. 5684238

; GENERAL INFORMATION:

; APPLICANT: Ausich, Rodney L.

; APPLICANT: Brinkhaus, Friedhelm L.

; APPLICANT: Mukharji, Indrani

; APPLICANT: Proffitt, John H.

; APPLICANT: Yarger, James G.

; APPLICANT: Yen, Hwei-Che B.

; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and

; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts

; NUMBER OF SEQUENCES: 104

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Welsh & Katz, Ltd.

; STREET: 120 S. Riverside Plaza, 22nd Floor

; CITY: Chicago

; STATE: IL

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/096,623A

; FILING DATE: 22-JUL-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/805,061

; FILING DATE: 09-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/662,921

; FILING DATE: 28-FEB-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/562,674

; FILING DATE: 03-AUG-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/525,551

; FILING DATE: 18-MAY-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/487,613

; FILING DATE: 02-MAR-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Gamson, Edward P.

; REGISTRATION NUMBER: 29,381

; REFERENCE/DOCKET NUMBER: AMO-006.1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (312) 655-1500

; TELEFAX: (312) 655-1501

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 177 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; POSITION IN GENOME:

; MAP POSITION: 1 to 177

; UNITS: bp

US-08-096-623A-12

Alignment Scores:

Pred. No.: 2,87e-28 Length: 177

Score: 254.00 Matches: 51

Percent Similarity: 96.6% Conservative: 6

Best Local Similarity: 86.4% Mismatches: 2

Query Match: 83.3% Indels: 0

DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x US-08-096-623A-12 (1-177)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCAGTCTCTTCTCTGAGCAGTTCGCCACCCGAGCAATGTTGCTCAAGCT 60

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AACATGGTGGCGCCTTTCACCTGCGCTTAAGTCAGTCGCTCATTCCTCTGTTTCAAGGAG 120

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59

Db 121 CAAAACCTTGACATCACTTCATTCGACGACGCGGAGAGTGCATGATGCAG 177

RESULT 10

PCT-US95-13937A-6

; Sequence 6, Application PC/TUS9513937A

; GENERAL INFORMATION:

; APPLICANT: Hauptmann, Randal

; APPLICANT: Eschenfeldt, William H

; APPLICANT: English, Jami

; APPLICANT: Brinkhaus, Friedhelm L

; TITLE OF INVENTION: Enhanced Carotenoid Accumulation

; TITLE OF INVENTION: in Storage Organs of Genetically

; TITLE OF INVENTION: Engineered Plants

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amoco Corporation, Law Dept

; STREET: 55 Shuman Boulevard, Suite 600

; CITY: Naperville

; STATE: IL

; COUNTRY: USA

; ZIP: 60563-8437

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/13937A

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Galloway, Norval B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 7087172447

; TELEFAX: 7087172430

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 177 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

PCT-US95-13937A-6

Alignment Scores:

Pred. No.: 2,87e-28 Length: 177

Score: 254.00 Matches: 51

Percent Similarity: 96.6% Conservative: 6

Best Local Similarity: 86.4% Mismatches: 2

Query Match: 83.3% Indels: 0

DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x PCT-US95-13937A-6 (1-177)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCAGTCTCTTCTCTGAGCAGTTCGCCACCCGAGCAATGTTGCTCAAGCT 60

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AACATGGTGGCGCTTCACTGGCCTTAAGTCAGCTGCCTCATTCCTCGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
|||||

Db 121 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATGCAG 177
|||||

RESULT 11

US-08-152-483B-8
; Sequence 8, Application US/08152483B
; Patent No. 5529909
; GENERAL INFORMATION:
; APPLICANT: Della-Cioppa, Guy
; APPLICANT: Kumagai, Monto
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
; TITLE OF INVENTION: PROTEIN FUSION ENZYME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,483B
; FILING DATE: No. 5529909ember 12, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7/857,602
; FILING DATE: March 30, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 923,692
; FILING DATE: July 31, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 600,244
; FILING DATE: October 22, 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 641,617
; FILING DATE: January 16, 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 737,899
; FILING DATE: July 26, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-3660
; TELEFAX: (415) 854-3694
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1442
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; IMMEDIATE SOURCE:
; CLONE:
; FEATURE:

US-08-152-483B-8

Alignment Scores:
Pred. No.: 3.18e-26 Length: 1442
Score: 249.00 Matches: 50
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.2% Mismatches: 2
Query Match: 81.6% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x US-08-152-483B-8 (1-1442)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||

Db 9 ATGGCTTCTCCTCAGTCTTTCTCTCTGCAGCAGTTGCCACCCGAGCAATGTTGCTCAAGCT 68
|||||

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
:::|||||

Db 69 AACATGGTTGCACCTTTCACCTGGCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 128
|||||

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMet 58
|||||

Db 129 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGAATGCATG 182
|||||

RESULT 12

US-09-441-340-9
; Sequence 9, Application US/09441340
; Patent No. 6448476
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Phosphonate Metabolizing Plants
; FILE REFERENCE: 38-21(15303)
; CURRENT APPLICATION NUMBER: US/09/441,340
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 60/108,763
; EARLIER FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:transit peptide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(264)
US-09-441-340-9

Alignment Scores:
Pred. No.: 2.58e-20 Length: 264
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: 3 Gaps: 1

US-10-718-311-15 (1-62) x US-09-441-340-9 (1-264)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||

Db 1 ATGGCTTCTCCTCAGTCTTTCTCTCTGCAGCAGTTGCCACCCGAGCAATGTTGCTCAAGCT 54
|||||

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
:::|||||

Db 55 ACTATGGTGGCTCTCTTCAACGGACTTAAGTCTCGCTGCCTTCCAGCCAGCCGCAAG 114
|||||

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
|||||

Db 115 GCTAACACGACATTAATTCTCCATCAAGCAACGGCGGAAGAGTTAACTGATGCAGGTG 174
|||||

Qy 61 Trp 61
|||

Db 175 TGG 177
|||

```
RESULT 13
US-09-186-002-5
; Sequence 5, Application US/09186002B
; Patent No. 6489542
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Romano, Charles P.
; TITLE OF INVENTION: Improved Method for Transforming Plants to Express
; TITLE OF INVENTION: delta-Endotoxins
; FILE REFERENCE: 38-21(13547) US Pat No. 6489542 09/186,002
; CURRENT APPLICATION NUMBER: US/09/186,002B
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; TYPE: DNA
; LENGTH: 268
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: transit_peptide
; LOCATION: (1)..(267)
; OTHER INFORMATION: coding sequence for PTPI comprising an Arabidopsis
; OTHER INFORMATION: thaliana ssRUBISCO (SSU) chloroplast targeting
; OTHER INFORMATION: sequence and sequences coding for the first 24
; OTHER INFORMATION: amino acids of ssRUBISCO (SSU) protein
US-09-186-002-5

Alignment Scores:
Pred. No.: 2,64e-20 Length: 268
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 10
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: 3 Gaps: 1

US-10-718-311-15 (1-62) x US-09-186-002-5 (1-268)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 54
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 55 ACTATGGTCGCTCTCTTCAACGAGCTTAAGTCTCCGCTGCCTTCCAGCCACCCGCAAG 114
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 115 GCTAACACAGCATTAATTCCATCACAGCAACGGCGGAGAGTTAACTGCATGCAGGTG 174
QY 61 Trp 61
Db 175 TGG 177

RESULT 14
US-08-391-339-9
; Sequence 9, Application US/08391339
; Patent No. 5463175
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr.,
; ADDRESS: Monsanto Co. B44F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

US-10-718-311-15 (1-62) x US-08-391-339-9 (1-279)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 12 ATGGCTTCCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 65
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 66 ACTATGGTCGCTCTCTTCAACGAGCTTAAGTCTCCGCTGCCTTCCAGCCACCCGCAAG 125
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 126 GCTAACACAGCATTAATTCCATCACAGCAACGGCGGAGAGTTAACTGCATGCAGGTG 185
QY 61 Trp 61
Db 186 TGG 188

RESULT 15
US-08-484-274A-9
; Sequence 9, Application US/08484274A
; Patent No. 5776760
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,339
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/156,968
; FILING DATE:
; APPLICATION NUMBER: US/07/717,370
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner, Dennis R., Jr.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10533)
; TELEPHONE: (314)537-6099
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant)
US-08-391-339-9
```

```
Alignment Scores:
Pred. No.: 2,8e-20 Length: 279
Score: 201.00 Matches: 41
Percent Similarity: 80.3% Conservative: 8
Best Local Similarity: 67.2% Mismatches: 10
Query Match: 65.9% Indels: 2
DB: 2 Gaps: 1

US-10-718-311-15 (1-62) x US-08-391-339-9 (1-279)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 12 ATGGCTTCCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 65
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 66 ACTATGGTCGCTCTCTTCAACGAGCTTAAGTCTCCGCTGCCTTCCAGCCACCCGCAAG 125
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 126 GCTAACACAGCATTAATTCCATCACAGCAACGGCGGAGAGTTAACTGCATGCAGGTG 185
QY 61 Trp 61
Db 186 TGG 188
```

```
RESULT 15
US-08-484-274A-9
; Sequence 9, Application US/08484274A
; Patent No. 5776760
; GENERAL INFORMATION:
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Barry, Gerard F.
; TITLE OF INVENTION: Glyphosate Tolerant Plants
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,274A
; FILING DATE: 07 June 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Patterson, Melinda L.
; REGISTRATION NUMBER: 33,062
; REFERENCE/DOCKET NUMBER: M08T130 38-21(13560)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713)789-2679
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant)
US-08-484-274A-9

Alignment Scores:
Pred. No.:      2.8e-20      Length:      279
Score:          201.00      Matches:      41
Percent Similarity: 80.3%      Conservative: 8
Best Local Similarity: 67.2%      Mismatches: 10
Query Match:      65.9%      Indels:      2
DB:               2          Gaps:      1

US-10-718-311-15 (1-62) x US-08-484-274A-9 (1-279)
Qy      1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db      12 ATGGCTTCTCTATGCTCTCTCCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 65
Qy      21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db      66 ACTATGGTCGCTCCCTTTCAACGGACTTAAGTCTCCGCTGCTTCCAGCCACCAGCAAG 125
Qy      41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db      126 GCTAAACAACGACATTACTTCCATCATCAAGCAACGCGGAGAGATTAACTGCATGCAGGTG 185
Qy      61 Trp 61
Db      186 TGG 188
```

Search completed: July 11, 2006, 21:02:57
Job time : 55.7059 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 19:48:48 ; Search time 1581.32 Seconds

(without alignments)
3288.703 Million cell updates/sec

Title: US-10-718-311-15

Perfect score: 305

Sequence: 1 MASSVISAANVATRSNVTOA.....LDITSIASNGRVSCMQVWH 62

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/abs/ABSSWEB spool/US10718311/runat_11072006_111523_5213/app.query.fasta_1
-DB=EST -OFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0 -L-LOOPCL=0 -L-LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs803p
-USER=US10718311@CN 1 1 12067 @runat_11072006_111523_5213 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_hic:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	297	97.4	205	BE462847	BE462847 EST325226
2	297	97.4	235	AW039464	AW039464 EST281745
3	297	97.4	237	AW039041	AW039041 EST281014
4	297	97.4	238	AW041024	AW041024 EST283888

5	297	97.4	245	1	AJ832092	AJ832092
6	297	97.4	250	7	AM442951	AM442951 EST307881
7	297	97.4	260	7	AW037699	AW037699 EST279328
8	297	97.4	260	7	AW094122	AW094122 EST287302
9	297	97.4	264	7	AW037518	AW037518 EST278774
10	297	97.4	264	7	AW037607	AW037607 EST279065
11	297	97.4	267	1	AI773328	AI773328 EST255028
12	297	97.4	277	1	AI773502	AI773502 EST254602
13	297	97.4	279	2	BI931188	BI931188 EST51077
14	297	97.4	279	7	BE462486	BE462486 EST324752
15	297	97.4	283	7	AW094091	AW094091 EST287271
16	297	97.4	286	7	BF051382	BF051382 EST436557
17	297	97.4	288	7	AW037811	AW037811 EST279440
18	297	97.4	294	7	AW040243	AW040243 EST282749
19	297	97.4	329	7	AW038470	AW038470 EST280153
20	297	97.4	330	7	AW092013	AW092013 EST285193
21	297	97.4	351	2	BI129648	BI129648 EST475294
22	297	97.4	356	2	BI927635	BI927635 EST547524
23	297	97.4	357	7	AW443393	AW443393 EST308323
24	297	97.4	367	7	AW039845	AW039845 EST282318
25	297	97.4	370	1	AI773548	AI773548 EST284648
26	297	97.4	372	7	BE463004	BE463004 EST325266
27	297	97.4	373	7	AW093805	AW093805 EST286985
28	297	97.4	375	2	BI123527	BI123527 EST469173
29	297	97.4	377	7	AW038793	AW038793 EST280749
30	297	97.4	382	1	AI777051	AI777051 EST252018
31	297	97.4	385	7	AW092154	AW092154 EST285250
32	297	97.4	393	7	AW040522	AW040522 EST283482
33	297	97.4	394	2	BI128649	BI128649 EST474295
34	297	97.4	395	7	AW217063	AW217063 EST295777
35	297	97.4	401	7	AW093581	AW093581 EST286761
36	297	97.4	404	2	BM535307	BM535307 EST588329
37	297	97.4	407	2	BI929701	BI929701 EST549590
38	297	97.4	408	7	AW038880	AW038880 EST280836
39	297	97.4	412	2	BI928909	BI928909 EST548798
40	297	97.4	416	7	AW094716	AW094716 EST287896
41	297	97.4	417	7	AW933080	AW933080 EST358923
42	297	97.4	420	7	AW041064	AW041064 EST283928
43	297	97.4	420	7	AW443412	AW443412 EST308342
44	297	97.4	421	2	BI133589	BI133589 EST466577
45	297	97.4	421	2	BG642684	BG642684 EST510878

ALIGNMENTS

RESULT 1
BE462847
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BE462847 205 bp mRNA linear EST 18-MAY-2001
EST325226 tomato flower buds 0-3 mm, Cornell University
Lycopersicon esculentum cDNA clone CTOA16011, mRNA sequence.
BE462847 1 GI:9508618
EST.
Lycopersicon esculentum (Solanum lycopersicum)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 205)
van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E.,
Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and
Tanksley, S.D.
Generation of ESTs from tomato flower tissue, 0-3 mm buds
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.
Location/Qualifiers
1..205

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOAL6Q11"
/tissue_type="flower"
/dev_stages="0-3mm buds"
/clone_lib="tomato flower buds 0-3 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Tankley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

ORIGIN

Alignment Scores:

Pred. No.:	2,096-30	Length:	205
Score:	297.00	Matches:	61
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	7	Gaps:	0

US-10-718-311-15 (1-62) x BE462847 (1-205)

```

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 ATGGCTTCTCTGTCATTTCTTCACAGCTGTTCACAGCAGCAGCAATGTTACACAGCT 67
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLyLys 40
Db 68 AGCATGGTGGACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 127
Qy 41 GlnAnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 128 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGTGCATGCAGGTG 187
Qy 61 Trp 61
Db 188 TGG 190

```

RESULT 2
AW039464
LOCUS
DEFINITION
EST281745 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET10112, mRNA sequence.

ACCESSION
AW039464
VERSION
AW039464.1 GI:5898218
KEYWORDS
EST.
SOURCE
Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM

REFERENCE
AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 235)

D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
Roning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,
Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
Giovannoni, J.

TITLE
JOURNAL
COMMENT
Generation of ESTs from tomato leaf tissue
Unpublished (1999)
Contact: CUGI

Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
1..235
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"

```

/clone="cLET10112"
/tissue_type="leaf"
/dev_stages="4-6 week old plants"
/lab_host="XLI-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX, EcoRI
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

```

ORIGIN

Alignment Scores:

Pred. No.:	2,516-30	Length:	235
Score:	297.00	Matches:	61
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	97.4%	Indels:	0
DB:	7	Gaps:	0

US-10-718-311-15 (1-62) x AW039464 (1-235)

```

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 14 ATGGCTTCTCTGTCATTTCTTCACAGCTGTTCACAGCAGCAGCAATGTTACACAGCT 73
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLyLys 40
Db 74 AGCATGGTGGACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 133
Qy 41 GlnAnLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 134 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGTGCATGCAGGTG 193
Qy 61 Trp 61
Db 194 TGG 196

```

RESULT 3

AW039041

LOCUS

DEFINITION

EST281014 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

clone cLET10C2, mRNA sequence.

ACCESSION

AW039041

VERSION

AW039041.1 GI:5897795

KEYWORDS

EST.

SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Eukaryota; Viridiplantae;

Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 237)

D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,

Roning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,

Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and

Giovannoni, J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

Contact: CUGI

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1..237

Location/Qualifiers

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="Rio Grande PtoR"

/db_xref="taxon:4081"

/clone="cLET10C2"

/tissue_type="leaf"

/dev_stage="4-6 week old plants"
/lab_host="Xll-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

ORIGIN

Alignment Scores:
Pred. No.: 2,54e-30 Length: 237
Score: 297.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.4% Indels: 0
DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW039041 (1-237)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||
Db 9 ATGGCTTCCTCTGCTATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 68
|||||
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
|||||
Db 69 AGCATGGTTGCACCTTTCATCTGCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 128
|||||
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
|||||
Db 129 CAAACCTTGACATCACTTCCATTTAGCAATGGTGGAGAGTTAGTGCATGCAGGTTG 188
|||||

QY 61 Trp 61
|||
Db 189 TGG 191

RESULT 4

AW041024

LOCUS

DEFINITION EST283888 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET615, mRNA sequence.

ACCESSION AW041024

VERSION AW041024.1 GI:5899778

KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 238)

REFERENCE

AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J.,
Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W.,
Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and
Giovannoni, J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1. .238
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET615"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="Xll-Blue MRF"

/clone_lib="tomato mixed elicitor, BTI"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

ORIGIN

Alignment Scores:
Pred. No.: 2,55e-30 Length: 238
Score: 297.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.4% Indels: 0
DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW041024 (1-238)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||
Db 19 ATGGCTTCCTCTGCTATTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 78
|||||
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
|||||
Db 79 AGCATGGTTGCACCTTTCATCTGCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 138
|||||
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
|||||
Db 139 CAAACCTTGACATCACTTCCATTTAGCAATGGTGGAGAGTTAGTGCATGCAGGTTG 198
|||||
QY 61 Trp 61
|||
Db 199 TGG 201

RESULT 5

AJ832092

LOCUS

DEFINITION

AJ832092 Lycopersicon esculentum var. cerasiforme fruit 12 dpa
Lycopersicon esculentum var. cerasiforme cDNA clone LEEA3_1, mRNA
sequence.

ACCESSION AJ832092

VERSION AJ832092.1 GI:53705017

KEYWORDS EST.

SOURCE

Lycopersicon esculentum var. cerasiforme (Solanum lycopersicum var.
cerasiforme)

ORGANISM

Lycopersicon esculentum var. cerasiforme
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 245)

REFERENCE

AUTHORS

Lemaire-Chamley, M., Petit, J., Garcia, V., Just, D., Baldet, P.,
Germain, V., Fagard, M., Moussier, M., Cheniclet, C. and Rothan, C.
Changes in transcriptional profiles are associated with early fruit
tissue specialization in tomato

Plant Physiol. 139 (2), 750-769 (2005)

16183847

Contact: Garcia V

Physiologie et Biotechnologie Vegetales

INRA

71 avenue E. Bourleaux, Villenave d'Ornon, 33883, FRANCE.

FEATURES

source

1. .245
Location/Qualifiers
/organism="Lycopersicon esculentum var. cerasiforme"
/mol_type="mRNA"
/cultivar="West Virginia 106"
/db_xref="taxon:195583"
/clone="LEEAA3_1"
/tissue_type="Fruit"
/dev_stage="12 dpa"
/clone_lib="Lycopersicon esculentum var. cerasiforme fruit
12 dpa"

ORIGIN

Alignment Scores:
 Pred. No.: 2,658-30 Length: 245
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 1 Gaps: 0

US-10-718-311-15 (1-62) x AJ832092 (1-245)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 Db 63 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGAGCAATGTTACACAAGCT 122
 Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 Db 123 AGCATGGTTGCACCTTTCACCTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 182
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db 183 CAAAACCTTGACATCATTCTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 242
 Qy 61 Trp 61
 Db 243 TGG 245

RESULT 6
 AW442951
 LOCUS
 DEFINITION EST307881 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 clone cLET42H11 5', mRNA sequence.
 ACCESSION AW442951
 VERSION AW442951.1 GI:6985133
 KEYWORDS
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
 Liang,F., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
 Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
 Giovannoni,J.J. and Martin,G.B.
 TITLE Generation of ESTs from tomato callus (mixed elicitor)
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 Location/Qualifiers
 1..250
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande PtoR"
 /db_xref="taxon:4081"
 /clone="cLET42H11"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisocotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

Pred. No.: 2,728-30 Length: 250
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW42951 (1-250)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 Db 6 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGAGCAATGTTACACAAGCT 65
 Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 Db 66 AGCATGGTTGCACCTTTCACCTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAAG 125
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 Db 126 CAAAACCTTGACATCATTCTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 185
 Qy 61 Trp 61
 Db 186 TGG 188

RESULT 7
 AW037699
 LOCUS
 DEFINITION EST279328 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 clone cLET4M17, mRNA sequence.
 ACCESSION AW037699
 VERSION AW037699.1 GI:5896453
 KEYWORDS
 SOURCE
 Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 260)
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
 Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
 Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
 Giovannoni,J.
 TITLE Generation of ESTs from tomato leaf tissue
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 Location/Qualifiers
 1..260
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande PtoR"
 /db_xref="taxon:4081"
 /clone="cLET4M17"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisocotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
 okadaic acid, or systemin prior to tissue harvest.
 site was destroyed during cloning."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,878-30 Length: 260
 Score: 297.00 Matches: 61

Percent Similarity: 100.0%
 Best Local Similarity: 100.0%
 Query Match: 97.4%
 DB: 7
 Indels: 0
 Gaps: 0

US-10-718-311-15 (1-62) x AW037699 (1-260)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 DB 12 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGACGAATGTTACACAAGCT 71
 |||||
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 |||||
 DB 72 AGCATGGTTGCACCTTCACTGCTCAAAATCTTCAGCCACTTTCCTGTTCAAAAGAAG 131
 |||||
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 DB 132 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 191
 |||||
 QY 61 Trp 61
 |||||
 DB 192 TGG 194

RESULT 8
 AW094122
 LOCUS EST287302 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION clone cLET27G20, mRNA sequence.

ACCESSION AW094122
 VERSION AW094122.1 GI:6059717

KEYWORDS EST.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 260)
 AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Niernan, W.,
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue
 JOURNAL Unpublished (1999)

COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source Location/Qualifiers

1..260
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande Ptor"
 /db_xref="taxon:4081"
 /clone="cLET27G20"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

ORIGIN

Alignment Scores:
 Pred. No.: 2,87e-30 Length: 260
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0%
 Best Local Similarity: 100.0%
 Query Match: 100.0%
 DB: 0
 Indels: 0
 Gaps: 0

Query Match: 97.4%
 DB: 7
 Indels: 0
 Gaps: 0

US-10-718-311-15 (1-62) x AW094122 (1-260)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 DB 6 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGACGACGAATGTTACACAAGCT 65
 |||||
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 |||||
 DB 66 ACATGGTTGCACCTTCACTGCTCAAAATCTTCAGCCACTTTCCTGTTCAAAAGAAG 125
 |||||
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 DB 126 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 185
 |||||
 QY 61 Trp 61
 |||||
 DB 186 TGG 188

RESULT 9

AW037518
 LOCUS EST278774 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION clone cLET4N17, mRNA sequence.

ACCESSION AW037518
 VERSION AW037518.1 GI:5896188

KEYWORDS EST.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 264)
 AUTHORS D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Niernan, W.,
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue
 JOURNAL Unpublished (1999)

COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source Location/Qualifiers

1..264
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande Ptor"
 /db_xref="taxon:4081"
 /clone="cLET4N17"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

ORIGIN

Alignment Scores:
 Pred. No.: 2,92e-30 Length: 264
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0%
 Best Local Similarity: 100.0%
 Query Match: 97.4%
 DB: 7
 Indels: 0
 Gaps: 0

US-10-718-311-15 (1-62) x AW037518 (1-264)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 12 ATGGCTTCTCTGTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 71
 |||||
 QY 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLyLys 40
 |||||
 Db 72 AGCATGGTGGACCTTTCAGTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 131
 |||||
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 132 CAARACCTTGACATCACTTCCATTGCTAGCAATGTTGGAGAGTTAGTGCATGCAAGTG 191
 |||||
 QY 61 Trp 61
 |||||
 Db 192 TGG 194

RESULT 10
 AW037607
 LOCUS EST279065 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION clone cLET4H5, mRNA sequence.
 ACCESSION AW037607
 VERSION AW037607.1 GI:5896361
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 264)
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
 Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
 Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
 Giovannoni,J.
 TITLE Generation of ESTs from tomato leaf tissue
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 Location/Qualifiers
 1..264

/organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande PtoR"
 /db_xref="taxon:4081"
 /clone="cLET4H5"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, etchylene, fenchion, EIX,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

ORIGIN

Alignment Scores:
 Pred. No.: 2,928-30 Length: 264
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW037607 (1-264)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 6 ATGGCTTCTCTGTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 65
 |||||
 QY 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLyLys 40
 |||||
 Db 66 AGCATGGTGGACCTTTCAGTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 125
 |||||
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 126 CAARACCTTGACATCACTTCCATTGCTAGCAATGTTGGAGAGTTAGTGCATGCAAGTG 185
 |||||
 QY 61 Trp 61
 |||||
 Db 186 TGG 188

RESULT 11
 AW037928
 LOCUS EST255028 tomato resistant, Cornell Lycopersicon esculentum CDNA
 DEFINITION clone cLER8L7, mRNA sequence.
 ACCESSION AW037928
 VERSION AW037928.1 GI:5271969
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 267)
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
 Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
 Giovannoni,J.J. and Martin,G.B.
 TITLE Generation of ESTs from Pseudomonas resistant tomato
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 Location/Qualifiers
 1..267

/organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="RII-12 (35S::Pto in Rio Grande x Money Maker)"
 /db_xref="taxon:4081"
 /clone="cLER8L7"
 /tissue_type="leaf"
 /dev_stage="4-week old"
 /lab_host="SOLR"
 /clone_lib="tomato resistant, Cornell"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cLER - Tomato Pseudomonas Resistant EST library.
 Directionally cloned cDNAs inserted into pBluescript
 SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN

Alignment Scores:
 Pred. No.: 2,97e-30 Length: 267
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 1 Gaps: 0

US-10-718-311-15 (1-62) x AW037928 (1-267)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 23 ATGGCTTCTCTGTGTCATTTCTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 82
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 DB 83 AGCATGGTTGCACCTTTCACCTGGCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 142
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 DB 143 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 202
 QY 61 TTP 61
 DB 203 TGG 205

RESULT 12
 LOCUS AI773502 277 bp mRNA linear EST 18-MAY-2001
 DEFINITION EST254602 tomato resistant, Cornell Lycopersicon esculentum cDNA
 clone cLER7A10, mRNA sequence.

ACCESSION AI773502
 VERSION AI773502.1 GI:5271543
 KEYWORDS Lycopersicon esculentum (Solanum lycopersicum)
 SOURCE Lycopersicon esculentum
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 277)
 AUTHORS D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
 Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
 Giovannoni,J.J. and Martin,G.B.

TITLE Generation of ESTs from Pseudomonas resistant tomato
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html

FEATURES
 source
 Location/Qualifiers
 1..277
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Ril-12 (35S::Pto in Rio Grande x Money Maker)"
 /db_xref="taxon:4081"
 /clone="CLER7A10"
 /tissue_type="leaf"
 /dev_stage="4-week old"
 /lab_host="SOLR"
 /clone_lib="tomato resistant, Cornell"

ORIGIN
 Alignment Scores:
 Pred. No.: 3,12e-30 Length: 277
 Score: 297.00 Matches: 61
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 97.4% Indels: 0
 DB: 1 Gaps: 0

US-10-718-311-15 (1-62) x AI773502 (1-277)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 DB 5 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTTCACACGCGAGCAATTTACACAAGCT 64
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 DB 65 AGCATGGTTGCACCTTTCACCTGGCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 124

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 DB 125 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 184
 QY 61 TTP 61
 DB 185 TGG 187

RESULT 13

LOCUS BI931188

DEFINITION

EST551077 tomato flower, 8 mm to preanthesis buds Lycopersicon
 esculentum cDNA clone cFOCI9C6 5' end, mRNA sequence.

ACCESSION BI931188

VERSION BI931188.1

KEYWORDS GI:16245660

SOURCE EST.

ORGANISM

Lycopersicon esculentum (Solanum lycopersicum)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 279)

AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
 Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
 Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -
 preanthesis

JOURNAL Unpublished (2001)

COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html

FEATURES

source

Location/Qualifiers

1..279

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="cFOCI9C6"

/tissue_type="flower"

/dev_stage="buds 8mm to preanthesis"

/clone_lib="tomato flower, 8 mm to preanthesis buds"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Cornell University; sequencing: The
 Institute for Genomic Research; Flower buds and flowers
 were taken from greenhouse plants (4-8 wks old, TA496).
 They were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

ORIGIN

Alignment Scores:

Pred. No.: 3,15e-30 Length: 279

Score: 297.00 Matches: 61

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 97.4% Indels: 0

DB: 2 Gaps: 0

US-10-718-311-15 (1-62) x BI931188 (1-279)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

DB 7 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTTCACACGCGAGCAATTTACACAAGCT 66

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

DB 67 AGCATGGTTGCACCTTTCACCTGGCTCAAAATCTTCAGCCACTTTCCTGTTACAAAGAAG 126

```

QY      41  GlnAsnLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
      |||
      127  CAAACCTTGACATCACTCCATTCCTAGCAATGGTGGAGAGTAGCTGCATGCAGGTG 186

QY      61  Trp 61
      |||
      187  TGG 189

RESULT 14
LOCUS   BE462486
DEFINITION EST324752 tomato flower buds 0-3 mm, Cornell University
          Lycopersicon esculentum cDNA clone cTOA1301, mRNA sequence.
ACCESSION BE462486
VERSION   BE462486.1  GI:9508255
KEYWORDS EST.
SOURCE   Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
          asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 279)
AUTHORS   van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E.,
          Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
          Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
          Tanksley,S.D.
TITLE     Generation of ESTs from tomato flower tissue, 0-3 mm buds
JOURNAL   Unpublished (1999)
COMMENT   Contact: CUGI
          Clemson University Genomics Institute
          100 Jordan Hall, Clemson, SC 29634, USA
          Email: http://www.genome.clemson.edu/orders/index.html
          5 prime sequence.
          Location/Qualifiers
            source
              1..279
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /cultivar="TA496"
                /db_xref="taxon:4081"
                /clone="cTOA1301"
                /tissue_type="Flower"
                /dev_stage="0-3mm buds"
                /clone_lib="tomato flower buds 0-3 mm, Cornell University"
                /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                XhoI; supplier: Tanksley; Flower buds and flowers were
                taken from greenhouse plants (4-8 wks old, TA496). They
                were immediately frozen in liquid nitrogen and then
                size-separated while remaining frozen."

ORIGIN
Alignment Scores:
Pred. No.: 3 15e-30 Length: 279
Score: 297.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.4% Indels: 0
DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x BE462486 (1-279)

QY      1  MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
      |||
      7  ATGGCTTCCTGTCATTCCTAGCAAGCTGTTCACACGACGCAATGTTACACAAGCT 66

QY      21  SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40
      |||
      67  AGCATGGTGGACCTTCCTAGCTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 126

QY      41  GlnAsnLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
      |||
      127  CAAACCTTGACATCACTCCATTCCTAGCAATGGTGGAGAGTAGCTGCATGCAGGTG 186

```

```

QY      61  Trp 61
      |||
      187  TGG 189

RESULT 15
LOCUS   AW094091
DEFINITION EST287271 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
          clone cLET27A10, mRNA sequence.
ACCESSION AW094091
VERSION   AW094091.1  GI:5059686
KEYWORDS EST.
SOURCE   Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
          asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 283)
AUTHORS   D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
          Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
          Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
          Giovannoni,J.
TITLE     Generation of ESTs from tomato leaf tissue
JOURNAL   Unpublished (1999)
COMMENT   Contact: CUGI
          Clemson University Genomics Institute
          100 Jordan Hall, Clemson, SC 29634, USA
          Email: http://www.genome.clemson.edu/orders/index.html
          5 prime sequence.
          Location/Qualifiers
            source
              1..283
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /cultivar="Rio Grande PtoR"
                /db_xref="taxon:4081"
                /clone="cLET27A10"
                /tissue_type="leaf"
                /dev_stage="4-6 week old plants"
                /lab_host="XLI-Blue MRP"
                /clone_lib="tomato mixed elicitor, BTI"
                /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
                XhoI; cLET - inoculated with a variety of disease response
                elicitors. Plants exposed to 2,6 dichloroisonicotinic
                acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
                okadaic acid, or systemin prior to tissue harvest. EcoRI
                site was destroyed during cloning."

ORIGIN
Alignment Scores:
Pred. No.: 3 21e-30 Length: 283
Score: 297.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.4% Indels: 0
DB: 7 Gaps: 0

US-10-718-311-15 (1-62) x AW094091 (1-283)

QY      1  MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
      |||
      12  ATGGCTTCCTGTCATTCCTAGCAAGCTGTTCACACGACGCAATGTTACACAAGCT 71

QY      21  SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40
      |||
      72  AGCATGGTGGACCTTTCACCTGGTCTCAAAATCTTCAGCCACTTCCCTGTTACAAAGAAG 131

QY      41  GlnAsnLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
      |||
      132  CAAACCTTGACATCACTTCATTCCTAGCAATGGTGGAGAGTAGCTGCATGCAGGTG 191

QY      61  Trp 61
      |||

```

Db 192 TGG 194

Search completed: July 11, 2006, 23:37:18
Job time : 1584.32 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 19:35:08 ; Search time 172.484 Seconds
(without alignments)
3759.291 Million cell updates/sec

Title: US-10-718-311-15
Perfect score: 305
Sequence: 1 MASSVSSAAVATRSNVTA.....LDITSIASNGRVSRCMQVWH 62

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10718311/runat_11072006_111518_5144/app_query.fasta.1
-DB=N Geneseg -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05p
-USER=US10718311 @CN 1.1 1423 @runat_11072006_111518_5144 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg 8:**

```
1: Geneseg1980s:*
2: Geneseg1990s:*
3: Geneseg2000s:*
4: Geneseg2001as:*
5: Geneseg2001bs:*
6: Geneseg2002as:*
7: Geneseg2002bs:*
8: Geneseg2003as:*
9: Geneseg2003bs:*
10: Geneseg2003cs:*
11: Geneseg2003ds:*
12: Geneseg2004as:*
13: Geneseg2004bs:*
14: Geneseg2005s:*
15: Geneseg2006s:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	305	100.0	684	6 ABA91838	Abag1838 Chloropla
2	305	100.0	684	6 ABK47712	Abk47712 DNA seque
3	305	100.0	684	10 ADC25988	Adc25988 Tomato/Es

4	305	100.0	684	14	ADV91622	Adv91622 Tomato ru
5	269	88.2	204	6	AAI70688	AAI70688 Tobacco R
6	269	88.2	297	10	ADK59826	Adk59826 Plant DNA
7	269	88.2	377	10	ADK56121	Adk56121 Plant DNA
8	269	88.2	489	10	ADK56120	Adk56120 Plant DNA
9	269	88.2	492	10	ADC76164	Adc76164 DNA homol
10	269	88.2	499	6	AAJ35635	AAJ35635 Nicotiana
11	269	88.2	504	10	ADK59824	Adk59824 Plant DNA
12	269	88.2	581	10	ADC75089	Adc75089 N bentham
13	269	88.2	608	10	ADC76944	Adc76944 DNA homol
14	269	88.2	614	10	ADC76956	Adc76956 DNA homol
15	269	88.2	615	10	ADC76949	Adc76949 DNA homol
16	269	88.2	632	10	ADC76953	Adc76953 DNA homol
17	269	88.2	684	10	ADC76948	Adc76948 DNA homol
18	269	88.2	718	10	ADK58382	Adk58382 Plant DNA
19	269	88.2	736	10	ADK54321	Adk54321 Plant DNA
20	269	88.2	736	10	ADK57660	Adk57660 Plant DNA
21	269	88.2	841	11	ADM44879	Adm44879 Insect re
22	269	88.2	847	11	ADM45447	Adm45447 Insect re
23	269	88.2	859	11	ADM45065	Adm45065 Insect re
24	269	88.2	924	10	ADC75566	Adc75566 DNA homol
25	269	88.2	958	10	ADC76165	Adc76165 DNA homol
26	269	86.9	619	11	ADM45158	Adm45158 Insect re
27	265	86.9	668	10	ADK59825	Adk59825 Plant DNA
28	265	86.9	668	11	ADM45687	Adm45687 Insect re
29	264	86.6	609	10	ADC76965	Adc76965 DNA homol
30	254	83.3	177	2	AAT27129	Aat27129 Modified
31	254	83.3	177	2	AAT40794	Aat40794 Chloropla
32	254	83.3	177	2	AAT37096	Aat37096 Ribulose
33	254	83.3	177	2	AAT41744	Aat41744 Tobacco t
34	254	83.3	177	2	AAT91545	Aat91545 Phytoene
35	250.5	82.1	683	10	ADK54330	Adk54330 Plant DNA
36	250.5	82.1	737	10	ADK57662	Adk57662 Plant DNA
37	250.5	82.1	737	10	ADK58383	Adk58383 Plant DNA
38	250.5	82.1	789	11	ADM44820	Adm44820 Insect re
39	250.5	82.1	806	11	ADM45686	Adm45686 Insect re
40	249	81.6	177	2	AAQ13721	Aaq13721 Phytoene
41	249	81.6	1442	2	AAQ2327	Aaq2327 Chloropla
42	244.5	80.2	655	11	ADM45154	Adm45154 Insect re
43	244.5	80.2	738	10	ADK54333	Adk54333 Plant DNA
44	244.5	80.2	753	10	ADK58381	Adk58381 Plant DNA
45	244.5	80.2	754	10	ADK57645	Adk57645 Plant DNA

ALIGNMENTS

RESULT 1

ABA91838

ID ABA91838 standard; DNA; 684 BP.

XX ABA91838;

XX 29-AUG-2003 (revised)

DT 15-MAY-2002 (first entry)

XX Chloroplast transit peptide-chorismate pyruvate lyase fusion gene.

XX Chloroplast transit peptide; tomato; Rubisco; plant;

KW ribulose-1,5-bisphosphate carboxylase; chorismate pyruvate lyase; CPL;

KW enzyme; p-hydroxybenzoic acid; transgenic plant; gene; ds.

XX Lycopersicon esculentum.

OS Escherichia coli.

OS Chimeric.

XX Key Location/Qualifiers

FT CDS 1..495

FT /*tag= a /product= "chloroplast-targeted CPL fusion"

FT transit_peptide 1..186

FT /*tag= b

FT /*note= "Tomato Rubisco chloroplast targeting sequence"

FT mat_peptide 187..492

/*tag= c
/note= "E. coli chorismate pyruvate lyase"

FT WO200194607-A2.
XX 13-DEC-2001.
XX 22-MAY-2001; 2001WO-US016661.
XX 02-JUN-2000; 2000US-0209854P.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX Meyer K, Van Dyk DE, Viitanen PV;
XX WPI; 2002-226795/28.
DR P-PSDB; AAM50959.
XX Producing para-hydroxy benzoic acid in green plant, comprises expression
PT of unique expression cassette containing gene encoding chorismate
PT pyruvate lyase operably linked to specific chloroplast targeting
PT sequence.
XX Claim 17; Page 57; 60pp; English.
XX The present nucleotide sequence is that of an open reading frame encoding
CC chloroplast-targeted chorismate pyruvate lyase (CPL) fusion protein TP-
CC CPL (see AAM50959). PCR (see ABA91841-42) was used to generate a DNA
CC fragment corresponding to the transit peptide of tomato Rubisco small
CC subunit and the first 4 amino acid residues of mature Rubisco. The DNA
CC fragment was ligated into pET24a-CPL, which carries the Escherichia coli
CC CPL open reading frame (see ABA91837). The TP-CPL construct is an example
CC of expression cassettes of the invention that are designed for the high-
CC level production of p-hydroxybenzoic acid (pHBA) in higher plants.
CC Cleavage of TP-CPL in the chloroplast releases a novel polypeptide (see
CC AAM50961) that has full enzyme activity, converting chorismate to pHBA. A
CC claimed plant comprising a CPL expression cassette is selected from
CC soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
CC barley, oats, sorghum, rice, Arabidopsis, sugarcane, canola,
CC millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric
CC component of liquid crystal polymers which have application in the
CC automotive, electrical and other industries. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,14e-33 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x ABA91838 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCTATTCTTACAGCTGTTCACAGCAGCAAGTGTACACAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGTGACCTTCACTGGTCTCAAAATCTTCAGCACCTTTCCTGTTACAAAGAAG 120
Qy 41 GlnAnLeuAapIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGTG 180
Qy 61 TrpHis 62
Db 181 TGGCAT 186

RESULT 2

ABK47712
ID ABK47712 standard; DNA; 684 BP.

XX AC ABK47712;
XX 18-JUN-2002 (first entry)
XX DNA sequence of ORF for TP-CPL fusion protein.

XX SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester; plant;
KW sinapoylglucose:malate sinapoyltransferase; glycosylated aromatic acid;
KW malate conjugated aromatic acid; polymer synthesis; ubiC; TP-CPL; gene;
KW carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;
KW tomato Rubisco small subunit precursor; rbcS2; mutant; ds.

XX Escherichia coli.
OS Lycopersicon esculentum.
OS Synthetic.
OS Chimeric.

XX Key Location/Qualifiers
FH CDS 1..684
FT /*tag= a
FT /product= "TP-CPL fusion protein"

XX WO200204653-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US021283.

XX 07-JUL-2000; 2000US-0216615P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Flint D, Meyer K, Viitanen PV;

XX WPI; 2002-303779/34.

XX P-PSDB; AAU77942.

XX Producing aromatic acid conjugates, involves contacting glycosylated
PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the
PT presence of sinapoylglucose:malate sinapoyltransferase.

XX Example; Page 70; 72pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana
CC gene (SNG1, sinapoylglucose accumulator 1) encoding
CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the
CC substitution of a glucose moiety on a glycosylated aromatic acid with a
CC malate moiety to form a malate conjugated aromatic acid. The enzyme is
CC useful for producing malate conjugated aromatic acids, carboxylic acid
CC conjugated aromatic acids or aromatic esters. The malate conjugated
CC aromatic acids are useful in the synthesis of various polymers. The
CC present sequence representing the ORF (open reading frame) for TP-CPL
CC (tomato Rubisco small subunit precursor for rbcS2-Escherichia coli
CC chorismate pyruvate lyase (CPL)) fusion protein is described in the
CC examples of the present invention

XX SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,14e-33 Length: 684
Score: 305.00 Matches: 62
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x ABK47712 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||

Db 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

Qy 61 TrpHis 62

Db 181 TGGCAT 186

RESULT 3

ID ADC25988 standard; DNA; 684 BP.

XX AC ADC25988;

XX 18-DEC-2003 (first entry)

XX Tomato/Escherichia coli chimeric TP-CPL DNA.

KW UDP-glucosyltransferase; pHBA; p-hydroxybenzoic acid ester glucoside;

KW liquid crystal polymer; LCP; methylparaben; preservative; food;

KW cosmetic industry; ds; gene; tomato; TP-CPL; chimeric;

KW chorismate pyruvate lyase; plant.

XX Chimeric.

OS Lycopersicon esculentum.

OS Escherichia coli.

XX Key Location/Qualifiers

FT CDS 1..684

FT /*tag= a

FT /product= "Tomato/Escherichia coli chimeric TP-CPL

FT protein"

XX WO2003066836-A2.

XX 14-AUG-2003.

XX 06-FEB-2003; 2003WO-US005863.

XX 07-FEB-2002; 2002US-0355511P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;

XX WPI; 2003-767259/72.

XX P-PSDB; ADC25989.

PT New nucleic acid encoding UDP-glucosyltransferase, useful for preparing

PT cells that produce p-hydroxybenzoic acid glucose ester, also the new

PT enzymes.

XX Example 9; SEQ ID NO 41; 161pp; English.

XX The invention relates to a novel isolated nucleic acid that encodes a UDP

CC -glucosyltransferase. The method of the invention may be used to

CC transform microorganisms or green plant cells so that these produce a

CC higher level of high-volume chemicals or materials, such as pHBA (p-

CC hydroxybenzoic acid) ester glucoside which is an intermediate for pHBA, a

CC monomer for liquid crystal polymers (LCPs) and starting material for

CC methylparaben. Methylparaben is a preservative commonly used in the food

CC and cosmetic industries. The encoded enzymes may be used for in vitro

CC production of these compounds and for identifying similar enzymes by

CC sequence comparison. The current sequence is that of the

CC tomato/Escherichia coli chimeric TP-CPL (chorismate pyruvate lyase) DNA

CC of the invention.

SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,14e-33 Length: 684

Score: 305.00 Matches: 62

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC25988 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 61 AGCATGGTTGCACCTTCACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTACAAAGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

Qy 61 TrpHis 62

Db 181 TGGCAT 186

RESULT 4

ID ADV91622 standard; DNA; 684 BP.

XX AC ADV91622;

XX 10-MAR-2005 (first entry)

XX Tomato rubisco small subunit precursor TP-E. coli CPL chimeric DNA.

KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;

KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;

KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;

KW genetically engineered microorganism; antioxidant; antimicrobial agent;

KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;

KW neoplasm; pHBA; para-hydroxybenzoic acid; p-hydroxybenzoic acid;

KW rubisco small subunit precursor; transit peptide;

KW chloroplast transit peptide; TP; chloroplast; gene; ds.

XX Lycopersicon esculentum.

OS Escherichia coli.

OS Chimeric.

XX Key Location/Qualifiers

FT CDS 1..684

FT /*tag= a

FT /product= "Tomato rubisco small subunit precursor TP-E.

FT coli CPL fusion protein"

XX US2004261147-A1.

XX 23-DEC-2004.

XX 16-JUN-2003; 2003US-00462162.

XX 16-JUN-2003; 2003US-00462162.

XX (MEYE/) MEYER K.

XX (VIIT/) VIITANEN P V.

XX (FLIN/) FLINT D.

XX Meyer K, Viitanen PV, Flint D;

XX WPI; 2005-057232/06.

XX P-PSDB; ADV91623.

XX Producing hydroquinone glucoside in a green plant comprises growing a
 PT green plant having nucleic acid fragments and chorismate pyruvate lyase
 PT expression cassette.
 XX
 XX
 XX Example 4; SEQ ID NO 33; 70pp; English.
 XX
 XX The present invention relates to methods and materials to produce
 CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)
 CC in genetically modified green plants and microorganisms. The method
 CC relies upon transgenic plants or genetically modified microorganisms that
 CC produce increased levels of the initial substrate para-hydroxybenzoic
 CC acid (p-hydroxybenzoic acid; pHA) in a biosynthetic pathway to produce
 CC arbutin. Plants and microbes can be genetically engineered to produce
 CC high levels of pHA either by functional expression of the bacterial
 CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-
 CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.
 CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-
 CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The
 CC present sequence is tomato rubisco small subunit precursor (rbcS2)
 CC transit peptide (chloroplast transit peptide; TP)-Escherichia coli CPL
 CC chimeric DNA. This sequence is present in the E. coli expression vector
 CC construct pET24a-TP-CPL.
 XX

Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5,148-33 Length: 684
 Score: 305.00 Matches: 62
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 14 Gaps: 0

US-10-718-311-15 (1-62) x ADV91622 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 1 ATGGCTTCTCTGTCTATTCTTCTCAGCAGCTGTTGCCACGACGCAATGTTACACAAGCT 60
 Qy 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40
 |||||
 Db 61 AGCATGGTGGACCTTTCAGTGGTCTCAAACTTCAGGCACCTTCCCTGTTACAAAGAAG 120
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAGAGTTAGTGCATGCGAGTG 180
 Qy 61 TrpHis 62
 |||||
 Db 181 TGGCAT 186

RESULT 5

AAI70688

ID AAI70688 standard; DNA; 204 BP.

XX

AC AAI70688;

XX

DT 04-FEB-2002 (first entry)

XX

DE Tobacco Rubisco transit peptide small subunit DNA.

XX

KW Tobacco; Rubisco; ribulosebiphosphate carboxylase; transit peptide;
 KW acetyl-CoA carboxylase; transgenic plant; oilseed; vegetable oil;
 KW Brassica; soybean; ds.

XX

OS Nicotiana tabacum.

XX

FH Key Location/Qualifiers

FT transit_peptide 1..171

FT /*tag= a

FT mat_peptide 172..204

FT /*tag= b

FT

FT /note= "5' end of mature protein coding sequence"
 XX
 XX WO200181604-A1.
 XX
 XX 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012928.
 XX
 XX 20-APR-2000; 2000US-0198794P.
 XX
 XX (CRGI) CARGILL INC.
 XX
 XX Shorosh BS, Debonte LR;
 XX
 XX WPI; 2002-041417/05.
 XX
 XX P-PSDB; AAM50331.
 XX
 XX Using Nucleic acids encoding alfalfa cytosolic Acetyl-Coenzyme A
 PT Carboxylase (ACCase; EC 6.4.1.2) to increase the oil content of Soybean
 PT and Brassica seeds.
 XX
 XX Disclosure; Fig 1; 62pp; English.
 XX

CC The present sequence is that of DNA encoding the tobacco small subunit
 CC Rubisco transit peptide and the 5' portion of the mature small subunit
 CC Rubisco protein (see AAM50331). Nucleic acid constructs of the invention
 CC include a nucleic acid encoding a cytosolic acetyl-CoA carboxylase (EC-
 CC 6.4.1.2) and a promoter, and may also include a nucleic acid encoding the
 CC tobacco small subunit Rubisco transit peptide. Such constructs can be
 CC introduced into soybean and Brassica plants (especially Brassica napus
 CC (oilseed rape), Brassica rapa (turnip), Brassica juncea, Brassica
 CC carinata, Brassica nigra (black mustard) and Brassica oleracea
 CC (cauliflower, sprout, cabbage, broccoli)) to increase the oil content of
 CC their seeds from about 5% to about 25% greater on a dry weight basis
 CC (claimed)
 XX

Sequence 204 BP; 50 A; 60 C; 43 G; 51 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,11e-28 Length: 204
 Score: 269.00 Matches: 53
 Percent Similarity: 96.7% Conservative: 6
 Best Local Similarity: 86.9% Mismatches: 2
 Query Match: 88.2% Indels: 0
 DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x AAI70688 (1-204)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 1 ATGGCTTCTCTCAGTCTTCTTCTCTGACAGCTTGCACCCGACGCAATGTTGCTCAAGCT 60
 Qy 21 SerMetValAlaProPheThrGlyLeuIlySerSerAlaThrPheProValThrIlyLys 40
 |||||
 Db 61 AACATGGTGGACCTTTCAGTGGCTTAAATGTCAGTGCCTCTCCCTGTTTCAAGGAAG 120
 Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAGAGTTAGTGCATGCGAGTG 180
 Qy 61 Trp 61
 |||||
 Db 181 TGG 183

RESULT 6

ADKS9826

ID ADKS9826 standard; DNA; 297 BP.

XX

AC ADKS9826;

XX

DT 06-MAY-2004 (first entry)

XX

DE Plant DNA sequence which confers altered metabolic characteristic #7209.

```

XX KW altered metabolic characteristic; plant; acid metabolism;
KW alcohol metabolism; fatty acid metabolism;
KW branched fatty acid metabolism; alkaloïd metabolism;
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX OS Unidentified.
XX DE WO2003020936-A1.
XX PN 13-MAR-2003.
XX PD 30-AUG-2002; 2002WO-US027884.
XX PF 31-AUG-2001; 2001US-0316471P.
XX PR (DOWC ) DOW CHEM CO.
XX PA (DOWC ) DOW AGROSCIENCES LLC.
XX PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX PI Oriado JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX XX WPI; 2003-313091/30.
XX XX Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX Claim 1; SEQ ID NO 7209; 2576pp; English.
XX XX The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloïd or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX SQ Sequence 297 BP; 83 A; 70 C; 70 G; 74 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.88e-28 Length: 297
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 18.2% Indels: 0
DB: Gaps: 0

US-10-718-311-15 (1-62) x ADK59826 (1-297)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 97 ATGGCTTCTCAGTCTTCTTCTCAGCAGCAGTTCGCCACCGCAGCAATGTGCTCAAGCT 156
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 157 AACATGGTTGCACCTTTCACAGCTCTTAAGCTGTGCTCCTCATTCCTCTTCAAGAAG 216
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 217 CAAACCTTGACATCACTTCCATTCGACGACGACGCGGGAAGAGTGAATGCATGCAGGTG 276
QY 61 Trp 61

```

```

Db 277 TGG 279
RESULT 7
ADK56121
ID ADK56121 standard; DNA; 377 BP.
XX ADK56121;
XX 06-MAY-2004 (first entry)
XX Plant DNA sequence which confers altered metabolic characteristic #3504.
XX altered metabolic characteristic; plant; acid metabolism;
XX alcohol metabolism; fatty acid metabolism;
XX branched fatty acid metabolism; alkaloïd metabolism;
XX amino acid metabolism; ester metabolism; glyceride metabolism;
XX phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX OS Unidentified.
XX PN WO2003020936-A1.
XX PD 13-MAR-2003.
XX PF 30-AUG-2002; 2002WO-US027884.
XX PR 31-AUG-2001; 2001US-0316471P.
XX PA (DOWC ) DOW CHEM CO.
XX PI (DOWC ) DOW AGROSCIENCES LLC.
XX PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX PI Oriado JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX XX WPI; 2003-313091/30.
XX XX Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX Claim 1; SEQ ID NO 3504; 2576pp; English.
XX XX The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloïd or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX SQ Sequence 377 BP; 103 A; 89 C; 86 G; 99 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.62e-28 Length: 377
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: Gaps: 0

US-10-718-311-15 (1-62) x ADK56121 (1-377)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

```

```
Db 68 ATGGCTTCTCAGTCTTCTCCTCAGCAGCAGTTCACCCGCGAGCAATGTTGCTCAAGCT 127
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaAlaThrPheProValThrLysLys 40
Db 128 AACATGGTTGACCTTTCACAGGCTCTTAAGTCTGCTGCTCATTCCTCTGTTTCAAGAAAG 187
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 198 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGCAATGCATGAGGTG 247
Qy 61 Trp 61
Db 248 TGG 250

RESULT 8
ADK56120
ID ADK56120 standard; DNA; 489 BP.
XX
AC ADK56120;
DT 06-MAY-2004 (first entry)
XX
DE Plant DNA sequence which confers altered metabolic characteristic #3503.
XX
KW altered metabolic characteristic; plant; acid metabolism;
KW alcohol metabolism; fatty acid metabolism;
KW branched fatty acid metabolism; alkaloid metabolism;
KW amino acid metabolism; ester metabolism; glyceride metabolism;
KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;
KW terpene metabolism; isoprenoid metabolism; alkene metabolism;
KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
OS Unidentified.
XX
PN WO2003020936-A1.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027884.
XX
PR 31-AUG-2001; 2001US-0316471P.
XX
PA (DOWC ) DOW CHEM CO.
XX
PA (DOWC ) DOW AGROSCIENCES LLC.
XX
PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
PI Orledo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX
DR WPI; 2003-313091/30.
XX
Novel genes that confer altered metabolic characteristics in Nicotiana
PT benhamiana plants, useful for altering the levels of metabolites e.g
PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
PS
Claim 1; SEQ ID NO 3503; 2576pp; English.
XX
The invention comprises DNA sequences which confer an altered metabolic
CC characteristic when they are expressed in a plant. The DNA sequences of
CC the invention are useful for producing plants with an altered metabolic
CC characteristic, such as: altered acid metabolism, alcohol metabolism,
CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
CC base metabolism, altered amino acid metabolism, altered ester metabolism,
CC altered glyceride metabolism, altered phenolic metabolism, altered
CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
CC metabolism, ketone or quinone metabolism. The DNA sequences of the
CC invention may be used to provide disease resistance in a plant and gene
CC shuffling or sexual PCR procedures. The present nucleic acid represents a
CC DNA sequence of the invention.
XX
SQ Sequence 489 BP; 136 A; 113 C; 114 G; 126 T; 0 U; 0 Other;
```

```
Alignment Scores: 3.77e-28 Length: 489
Pred. No.: 269.00 Matches: 53
Score: 96.7% Conservative: 6
Percent Similarity: 96.7% Mismatches: 2
Best Local Similarity: 86.9% Indels: 0
Query Match: 88.2% Gaps: 0
DB: 10
US-10-718-311-15 (1-62) x ADK56120 (1-489)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 79 ATGGCTTCTCAGTCTTCTCCTCAGCAGCAGTTCGCCACCCGCGAGCAATGTTGCTCAAGCT 138
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaAlaThrPheProValThrLysLys 40
Db 139 AACATGGTTGACCTTTCACAGGCTCTTAAGTCTGCTGCTCATTCCTCTGTTTCAAGAAAG 198
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 199 CAAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAGAGTGCATGCAATGCATGAGGTG 258
Qy 61 Trp 61
Db 259 TGG 261

RESULT 9
ADK76164
ID ADC76164 standard; DNA; 492 BP.
XX
AC ADC76164;
XX
DT 01-JAN-2004 (first entry)
XX
DE DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1433.
XX
KW rice; yeast; poppy; plant; disease resistance; anti-fungal;
KW phytopathogen; gene shuffling; ds.
XX
OS Unidentified.
XX
PN WO2003020905-A2.
XX
PD 13-MAR-2003.
XX
PF 30-AUG-2002; 2002WO-US027883.
XX
PR 31-AUG-2001; 2001US-0316392P.
XX
PA (DOWC ) DOW CHEM CO.
XX
PI Shukla V, Butler H, Larrinua I, Reddy AS;
XX
DR WPI; 2003-290185/28.
XX
Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryza
PT sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
PT rhoeas, useful for conferring disease resistance in plants.
PS
Claim 1; SEQ ID NO 1433; 617pp; English.
XX
The invention relates to a novel isolated nucleic acid derived from
CC Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
CC (yeast), Trichoderma harzianum (Hypocrea lixi) and Papaver rhoeas
CC (poppy), and a sequence that hybridises to them under conditions of low
CC stringency, where expression of the nucleic acid in a plant results in a
CC disease resistance phenotype. The polynucleotides of the invention
CC demonstrate anti-fungal activity and may be useful in conferring disease
CC resistance in a plant against phytopathogen such as Aspergillus flavus,
CC Gibberella fujikuroi and Gibberella zeae. Furthermore, the
CC polynucleotides may be useful to retrieve unknown sequences and in gene
CC shuffling or sexual PCR procedures. The current sequence is that of the
```

CC DNA of the invention which is homologous to that of the phytopathogen
CC resistance-related contig cDNAs.

XX Sequence 492 BP; 140 A; 115 C; 111 G; 126 T; 0 U; 0 Other;

SQ

Alignment Scores:
Pred. No.: 3.8e-28 Length: 492
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC76164 (1-492)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 66 ATGGCTTCTCAGTTCTTCTCCTCAGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 125
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 126 AACATGGTTGCACCTTTCTACCTGCTCTTAAGTCTGCTGCCTCATTCCTCTTTCAAGGAAG 185
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 186 CAAACCTTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 245
QY 61 Ttp 61
Db 246 TGG 248

RESULT 10

AAD35635
ID AAD35635 standard; cDNA; 499 BP.

XX AC AAD35635;

XX DT 26-JUL-2002 (first entry)

XX DE Nicotiana benthamiana rubisco partial cDNA sequence.

XX KW Nucleic acid vector; tobacco rattle virus; virus induced gene silencing;
KW TRV; VIGS; ribulose-1,5-bisphosphate carboxylase oxygenase; rubisco;
KW transgenic plant; enzyme; ss.

XX OS Nicotiana benthamiana.

XX PN US6369296-B1.

XX PD 09-APR-2002.

XX PF 01-FEB-2000; 2000US-00495797.

XX PR 01-FEB-2000; 2000US-00495797.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX PI Ratcliff FG, Martin-Hernandez AM, Baulcombe DC;

XX DR WPI; 2002-337969/37.

XX PT Novel nucleic acid vector, useful for producing transgenic plants,
PT comprises a plant active promoter linked to a recombinant tobacco rattle
PT virus cDNA.

XX PS Example 2; Col 37-38; 32pp; English.

XX CC The invention relates to a nucleic acid vector. The vector comprises a
CC plant active promoter operably linked to a recombinant tobacco rattle
CC virus (TRV) cDNA which includes cis acting elements of TRV RNA permitting
CC in the presence of replicase, replication of the cDNA, a viral subgenomic
CC promoter operably linked to a sequence encoding a TRV coat protein, and a
CC heterologous sequence, and border sequences for transfer into a plant

CC genome. The vector is useful for producing genetically engineered plant
CC cells or transgenic plants. The nucleic acids are used to initiate
CC virus induced gene silencing (VIGS). The present sequence is Nicotiana
CC benthamiana ribulose-1,5-bisphosphate carboxylase oxygenase (rubisco)
CC partial cDNA related to the invention

SQ Sequence 499 BP; 132 A; 120 C; 123 G; 124 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.88e-28 Length: 499
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 6 Gaps: 0

US-10-718-311-15 (1-62) x AAD35635 (1-499)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 ATGGCTTCTCAGTTCTTCTCCTCAGCAGTTCGCCACCGCAGCAATGTTGCTCAAGCT 67
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 68 ACATGGTTGCACCTTTCTACCTTGAAGTACGTGCTCATTCCTCTTTCAAGGAAG 127
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAACCTTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCAATGCATGCAGGTG 187
QY 61 Ttp 61
Db 188 TGG 190

RESULT 11

ADK59824

ID ADK59824 standard; DNA; 504 BP.

XX AC ADK59824;

XX DT 06-MAY-2004 (first entry)

XX DE Plant DNA sequence which confers altered metabolic characteristic #7207.

XX KW altered metabolic characteristic; plant; acid metabolism;

XX KW alcohol metabolism; fatty acid metabolism;

XX KW branched fatty acid metabolism; alkaloid metabolism;

XX KW amino acid metabolism; ester metabolism; glyceride metabolism;

XX KW phenolic metabolism; carbohydrate metabolism; sterol metabolism;

XX KW terpene metabolism; isoprenoid metabolism; alkene metabolism;

XX KW alkyne metabolism; hydrocarbon metabolism; ketone metabolism;

XX KW quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.

XX OS Unidentified.

XX PN WO2003020936-A1.

XX PD 13-MAR-2003.

XX PF 30-AUG-2002; 2002WO-US027884.

XX PR 31-AUG-2001; 2001US-0316471P.

XX PA (DOWC) DOW CHEM CO.

XX PA (DOWC) DOW AGROSCIENCES LLC.

XX PI Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;

XX PI Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;

XX WPI; 2003-313091/30.

XX PT Novel genes that confer altered metabolic characteristics in Nicotiana
PT benthamiana plants, useful for altering the levels of metabolites e.g.

PT acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
 XX
 PS Claim 1; SEQ ID NO 7207; 2576pp; English.
 XX
 CC The invention comprises DNA sequences which confer an altered metabolic
 CC characteristic when they are expressed in a plant. The DNA sequences of
 CC the invention are useful for producing plants with an altered metabolic
 CC characteristic, such as: altered acid metabolism, alcohol metabolism,
 CC fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
 CC base metabolism, altered amino acid metabolism, altered ester metabolism,
 CC altered glyceride metabolism, altered phenolic metabolism, altered
 CC carbohydrate metabolism, altered sterol, oxygenated terpene, or
 CC isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
 CC metabolism, ketone or quinone metabolism. The DNA sequences of the
 CC invention may be used to provide disease resistance in a plant and gene
 CC shuffling or sexual PCR procedures. The present nucleic acid represents a
 CC DNA sequence of the invention.
 XX
 SQ Sequence 504 BP; 139 A; 117 C; 118 G; 130 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.93e-28 Length: 504
 Score: 269.00 Matches: 53
 Percent Similarity: 96.7% Conservative: 6
 Best Local Similarity: 86.9% Mismatches: 2
 Query Match: 88.2% Indels: 0
 DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADK59824 (1-504)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 78 ATGGCTTCTCAGTCTTCTCTCAGCAGCAGTTCGCCACCGCAGCAATGTTGCTCAAGCT 137
 |||||
 Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 ::::|
 Db 138 AACATGGTTGCACCTTTCACAGGTCTTAAGTCTGCTCCTCATTCCTCTGTTTCAAGAAAG 197
 |||||
 Qy 41 GlnAnLeuAspIleThrSerIleAlaSerAnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 198 CAAACCTTGCATCACTTCCATTGCCAGCAACGCGGAAGAGTGCATGCAATGCAGGTG 257
 |||||
 Qy 61 Trp 61
 |||||
 Db 258 TGG 260

RESULT 12
 ADC75089
 ID ADC75089 standard; cDNA; 581 BP.
 XX
 AC ADC75089;

DT 01-JAN-2004 (first entry)

XX
 DE N benthamiana phytopathogen resistance-related contig cDNA - SEQ ID 13.
 XX
 KW rice; yeast; poppy; plant; disease resistance; anti-fungal;
 KW phytopathogen; gene shuffling; ss.

XX Nicotiana benthamiana.

OS WO2003020905-A2.

PN 13-MAR-2003.

XX 30-AUG-2002; 2002WO-US027883.

XX 31-AUG-2001; 2001US-0316392P.

XX (DOWC) DOW CHEM CO.

XX Shukla V, Butler H, Larrinua I, Reddy AS;

XX

DR WPI; 2003-290185/28.

XX
 PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
 PT sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
 PT rhoeas, useful for conferring disease resistance in plants.
 XX

PS Claim 1; SEQ ID NO 13; 617pp; English.

XX The invention relates to a novel isolated nucleic acid derived from
 CC Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
 CC (yeast), Trichoderma harzianum (Hypocrea lixi) and Papaver rhoeas
 CC (poppy) and a sequence that hybridises to them under conditions of low
 CC stringency, where expression of the nucleic acid in a plant results in a
 CC disease resistance phenotype. The polynucleotides of the invention
 CC demonstrate anti-fungal activity and may be useful in conferring disease
 CC resistance in a plant against phytopathogen such as Aspergillus flavus,
 CC Gibberella fujikuroi and Gibberella zeae. Furthermore, the
 CC polynucleotides may be useful to retrieve unknown sequences and in gene
 CC shuffling or sexual PCR procedures. The current sequence is that of the
 CC phytopathogen resistance-related contig cDNA of the invention.
 XX

SQ Sequence 581 BP; 162 A; 132 C; 143 G; 144 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.8e-28 Length: 581
 Score: 269.00 Matches: 53
 Percent Similarity: 96.7% Conservative: 6
 Best Local Similarity: 86.9% Mismatches: 2
 Query Match: 88.2% Indels: 0
 DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC75089 (1-581)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 63 ATGGCTTCTCAGTCTTCTCTCAGCAGCAGTTCGCCACCGCAGCAATGTTGCTCAAGCT 122
 |||||
 Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 ::::|
 Db 123 AACATGGTTGCACCTTTCACAGGTCTTAAGTCTGCTCCTCATTCCTCTGTTTCAAGAAAG 182
 |||||
 Qy 41 GlnAnLeuAspIleThrSerIleAlaSerAnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 183 CAAACCTTGCATCACTTCCATTGCCAGCAACGCGGAAGAGTGCATGCAATGCAGGTG 242
 |||||
 Qy 61 Trp 61
 |||||
 Db 243 TGG 245

RESULT 13
 ADC76944
 ID ADC76944 standard; DNA; 608 BP.
 XX
 AC ADC76944;

DT 01-JAN-2004 (first entry)

XX DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1213.
 XX
 KW rice; yeast; poppy; plant; disease resistance; anti-fungal;
 KW phytopathogen; gene shuffling; ds.

XX Unidentified.

PN WO2003020905-A2.

XX 13-MAR-2003.

XX 30-AUG-2002; 2002WO-US027883.

XX 31-AUG-2001; 2001US-0316392P.

XX (DOWC) DOW CHEM CO.


```

XX  Shukla V, Butler H, Larrinua I, Reddy AS;
PI  WPI; 2003-290185/28.
XX
XX  Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
PT  sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
PT  rhoeas, useful for conferring disease resistance in plants.
XX
XX  Claim 1; SEQ ID NO 1213; 617pp; English.
XX
XX  The invention relates to a novel isolated nucleic acid derived from
CC  Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
CC  (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoeas
CC  (poppy) and a sequence that hybridises to them under conditions of low
CC  stringency, where expression of the nucleic acid in a plant results in a
CC  disease resistance phenotype. The polynucleotides of the invention
CC  demonstrate anti-fungal activity and may be useful in conferring disease
CC  resistance in a plant against phytopathogen such as Aspergillus flavus,
CC  Gibberella fujikuroi and Gibberella zeae. Furthermore, the
CC  polynucleotides may be useful to retrieve unknown sequences and in gene
CC  shuffling or sexual PCR procedures. The current sequence is that of the
CC  DNA of the invention which is homologous to that of the phytopathogen
CC  resistance-related contig cDNAs.
XX
SQ  Sequence 608 BP; 164 A; 144 C; 142 G; 158 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,11e-28 Length: 608
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC76944 (1-608)

QY  1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB  38 ATGGCTTCCTCAGTTCTTCTCAGCAGCAGTGGCCACCGCAGCAATGTGCTCAAGCT 97
QY  21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB  98 AACATGGTGGCACCCTTTCACAGGTCTTAAGTCGTGCTCCTTCCTGTTTCAAGAAAG 157
QY  41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB  158 CAAACCTTGACATCACTTCCATTCGACGACGCGGAGAGTGCATGCAATGCATGCAGGTG 217
QY  61 Trp 61
DB  218 TGG 220

RESULT 14
ADC76956
ID  ADC76956 standard; DNA; 614 BP.
XX
XX  ADC76956;
AC
AC  01-JAN-2004 (first entry)
DT
DE  DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1225.
XX
XX  rice; yeast; poppy; plant; disease resistance; anti-fungal;
KW  phytopathogen; gene shuffling; ds.
XX
XX  Unidentified.
OS
XX  WO2003020905-A2.
XX
XX  13-MAR-2003.
XX
XX  30-AUG-2002; 2002WO-US027883.
PF

```

```

XX  31-AUG-2001; 2001US-0316392P.
XX
XX  (IDMC ) DOW CHEM CO.
XX
XX  Shukla V, Butler H, Larrinua I, Reddy AS;
PI  WPI; 2003-290185/28.
XX
XX  Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
PT  sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
PT  rhoeas, useful for conferring disease resistance in plants.
XX
XX  Claim 1; SEQ ID NO 1225; 617pp; English.
XX
XX  The invention relates to a novel isolated nucleic acid derived from
CC  Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
CC  (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoeas
CC  (poppy) and a sequence that hybridises to them under conditions of low
CC  stringency, where expression of the nucleic acid in a plant results in a
CC  disease resistance phenotype. The polynucleotides of the invention
CC  demonstrate anti-fungal activity and may be useful in conferring disease
CC  resistance in a plant against phytopathogen such as Aspergillus flavus,
CC  Gibberella fujikuroi and Gibberella zeae. Furthermore, the
CC  polynucleotides may be useful to retrieve unknown sequences and in gene
CC  shuffling or sexual PCR procedures. The current sequence is that of the
CC  DNA of the invention which is homologous to that of the phytopathogen
CC  resistance-related contig cDNAs.
XX
SQ  Sequence 614 BP; 167 A; 145 C; 149 G; 153 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,18e-28 Length: 614
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC76956 (1-614)

QY  1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB  46 ATGGCTTCCTCAGTTCTTCTCAGCAGCAGTGGCCACCGCAGCAATGTGCTCAAGCT 105
QY  21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB  106 AACATGGTGGCACCCTTTCACAGGTCTTAAGTCAGTGCCTTCCTGTTTCAAGGAAG 165
QY  41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB  166 CAAACCTTGACATCACTTCCATTCGACGACGCGGAGAGTGCATGCAATGCATGCAGGTG 225
QY  61 Trp 61
DB  226 TGG 228

RESULT 15
ADC76949
ID  ADC76949 standard; DNA; 615 BP.
XX
XX  ADC76949;
AC
AC  01-JAN-2004 (first entry)
DT
DE  DNA homologous to phytopathogen resistance-related cDNA - SEQ ID 1218.
XX
XX  rice; yeast; poppy; plant; disease resistance; anti-fungal;
KW  phytopathogen; gene shuffling; ds.
XX
XX  Unidentified.
OS
XX  WO2003020905-A2.
PN

```

```

XX 13-MAR-2003.
XX 30-AUG-2002; 2002MO-US027883.
XX 31-AUG-2001; 2001US-0316392P.
XX (DOWC ) DOW CHEM CO.
XX Shukla V, Butler H, Larrinua I, Reddy AS;
XX WPI; 2003-290185/28.
XX Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
XX sativa, Saccharomyces cerevisiae, Trichoderma harzianum and Papaver
XX rhoeas, useful for conferring disease resistance in plants.
XX Claim 1; SEQ ID NO 1218; 617pp; English.
XX The invention relates to a novel isolated nucleic acid derived from
XX Nicotiana benthamiana, Oryza sativa (rice), Saccharomyces cerevisiae
XX (yeast), Trichoderma harzianum (Hypocrea lixii) and Papaver rhoeas
XX (poppy) and a sequence that hybridises to them under conditions of low
XX stringency, where expression of the nucleic acid in a plant results in a
XX disease resistance phenotype. The polynucleotides of the invention
XX demonstrate anti-fungal activity and may be useful in conferring disease
XX resistance in a plant against phytopathogen such as Aspergillus flavus,
XX Gibberella fujikuroi and Gibberella zeae. Furthermore, the
XX polynucleotides may be useful to retrieve unknown sequences and in gene
XX shuffling or sexual PCR procedures. The current sequence is that of the
XX DNA of the invention which is homologous to that of the phytopathogen
XX resistance-related contig cDNAs.
XX SQ Sequence 615 BP; 168 A; 144 C; 146 G; 156 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 5.2e-28 Length: 615
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 88.2% Indels: 0
DB: 10 Gaps: 0

US-10-718-311-15 (1-62) x ADC76949 (1-615)
QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 47 ATGGCTTCCTCAGTTCTTTCCAGCAGCAGTTCGCCGCCGAGCAATGTTGCTCAAGCT 106
QY 21 SerMetValAlaProPheThrGlyLeuLySerSerAlaThrPheProValThrLysLys 40
Db 107 AACATGGTTGCACCTTTCACAGCTCTTAAGTCTGCTGCTCATTCCTCTTCAAGAAAG 166
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 167 CAAAACCTTGACATCACTTCCATTCGACGACACGGCGGAGAGAGTCAATGCATGCAGGTG 226
QY 61 Trp 61
Db 227 TGG 229

Search completed: July 11, 2006, 21:50:41
Job time : 173.484 secs

```

```

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 20:58:40 ; Search time 135.1 Seconds
(without alignments)
3276.788 Million cell updates/sec

Title: US-10-718-311-8
Sequence: 1 MASSVIVSAVATRSNVTOA.....RLSGKPLLTFLPASPLY 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 844069 seqs, 650066433 residues
Total number of hits satisfying chosen parameters: 1688138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abes/ABSWEB/spool/US10718311/runat_11072006_111535_5434/app_query.fasta.1
-DB=Published Applications NA New -QWMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -HOST=abs06p
-USER=US10718311 @CGN 1.1.251 @runat_11072006_111535_5434 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description
-----
1 281.5 24.3 781 8 US-11-247-587-124 Sequence 124, App
2 276.5 23.9 782 8 US-11-247-587-125 Sequence 125, App
3 253.5 21.9 698 8 US-11-247-587-128 Sequence 128, App
4 247.5 21.4 683 8 US-11-247-587-126 Sequence 126, App
5 246.5 21.3 700 8 US-11-247-587-127 Sequence 127, App
6 238.5 20.6 583 7 US-11-218-305-17631 Sequence 17631, A
7 231.5 20.0 727 8 US-11-247-587-129 Sequence 129, App
8 228.5 19.7 743 6 US-10-953-349-8894 Sequence 8894, Ap

Sequence 17633, A
Sequence 10521, A
Sequence 17634, A
Sequence 4173, Ap
Sequence 9569, Ap
Sequence 17632, A
Sequence 4174, Ap
Sequence 9381, Ap
Sequence 36894,
Sequence 388679,
Sequence 450273,
Sequence 24132, A
Sequence 6801, Ap
Sequence 17846, A
Sequence 1748, A
Sequence 20231, A
Sequence 18147, A
Sequence 17, Appl
Sequence 13, Appl
Sequence 8, Appl
Sequence 6032, Ap
Sequence 368279,
Sequence 451658,
Sequence 32248, A
Sequence 17356, A
Sequence 14808, A
Sequence 6036, Ap
Sequence 1067, Ap
Sequence 28461, A
Sequence 12983, A
Sequence 295512,
Sequence 346941,
Sequence 239, App
Sequence 18012, A
Sequence 14286, A
Sequence 12992, A

```

ALIGNMENTS

```

RESULT 1
US-11-247-587-124
; Sequence 124, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 124
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-124
Alignment Scores:
Pred. No.: 1.39e-22 Length: 781
Score: 281.50 Matches: 80
Percent Similarity: 58.3%
Conservative: 22

```


DB: 8 Caps: 8
US-10-718-311-8 (1-227) x US-11-247-587-128 (1-698)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
DB 20 ATGGCTTCCTCTGTGATTCCTCAGCTGTGCGGTTGCCACCGCGCTAATGCTGCTCAA 79
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39
DB 80 GCAGCATGTTGCACCCCTTCACTGGCCTCAAAATCTGCTTCCTCTCCCTGTTACCAGA 139
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
DB 140 AAACAAACCTTGACATTCATCCATGCTAGCAATGGTGGAGAGTCCAAATGCATGCAG 199
QY 60 ValTTPHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGlu 79
DB 200 GTGTGGCCCAATTAAC-----ATGAAGAAGTACGACACTCTCATACCTTCCTGAT 253
QY 80 IleProAlaLeuAspProGlnLeuLeuAsp-----TTP 90
DB 254 TTGAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTATCTTTTGAATAATGATGG 307
QY 91 LeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThrValSerVal 110
DB 308 GTTCCTTGTGGAATTCGAGACTGAGCGT-----337
QY 111 ThrMetIleArgGluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuPro 128
DB 338 -----GGATTGCTACCGTGAACATCAGCTACAGGATGACTACTAC---379
QY 129 LeuLeuProLysGluSerArgTyrTrpLeuArg-----GluLeuLeuCysAla 145
DB 380 -----GATGGCAGATACGACCATGTGAAGTTGGCCACTGTTCGGGTGCACT 427
QY 146 AspGlyGluProTrpLeuAla 152
DB 428 GATGCCACTCAGGTGTGGCT 448

RESULT 4

US-11-247-587-126
; Sequence 126, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 126
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-126

Alignment Scores:
Pred. No.: 8,14e-19 Length: 683
Score: 247.50 Matches: 70
Percent Similarity: 52.7% Conservative: 18
Best Local Similarity: 41.9% Mismatches: 40

Query Match: 21.4% Indels: 39
DB: 8 Gaps: 8
US-10-718-311-8 (1-227) x US-11-247-587-126 (1-683)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
DB 37 ATGGCTTCCTCTGTGATTCCTCAGCAGCTGTGTTGCGCGCGCCCAATGCTGCTCAA 96
QY 20 AlaSerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLys 39
DB 97 GCCAATGTTGGCAGCCCTTCACTGGCCTCAAGTCGCTCCTCTCCCTGTTACCAG 156
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
DB 157 AAACAAACCTTGACATTCATCCATGCTAGCAATGGTGGAGAGTCCAAATGCATGCAG 216
QY 60 ValTTPHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGlu 79
DB 217 GTGTGGCCCAATTAAC-----ATGAAGAAGTACGACACTCTCATACCTTCCTGAT 270
QY 80 IleProAlaLeuAspProGlnLeuLeuAsp-----TTP 90
DB 271 TTGAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTATCTTTTGAATAATGATGG 324
QY 91 LeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThrValSerVal 110
DB 325 GTTCCTTGTGGAATTCGAGACTGAGCGT-----354
QY 111 ThrMetIleArgGluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuPro 128
DB 355 -----GGATTGCTACCGTGAACATCAGCTACAGGATGACTACTAC---396
QY 129 LeuLeuProLysGluSerArgTyrTrpLeuArg-----GluLeuLeuCysAla 145
DB 397 -----GATGGCAGATACGACCATGTGAAGTTGGCCACTGTTCGGGTGCACT 444
QY 146 AspGlyGluProTrpLeuAla 152
DB 445 GATGCCACTCAGGTGTGGCT 465

RESULT 5

US-11-247-587-127
; Sequence 127, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-127

Alignment Scores:
Pred. No.: 1.1e-18 Length: 700
Score: 246.50 Matches: 71
Percent Similarity: 52.1% Conservative: 17

```
Best Local Similarity: 42.0% Mismatches: 38
Query Match: 21.3% Indels: 43
DB: 8 Gaps: 9

US-10-718-311-8 (1-227) x US-11-247-587-127 (1-700)
QY 1 MetAlaSerSerValIleSerSer---AlaAlaValAlaThrArgSerAsnValThrGln 19
Db 2 ATGGCTTCTCAGTTATGCTCAGCTGCGCTGTTGCCACGGCGCAATGCTCAA 61
QY 20 AlaSerMetValAlaProPheThrGlyLeuIleYssSerAlaThrPheProValThrLys 39
Db 62 GCCAGTATGGTGCACCTTCACTGGCCTCAAGTCGCAACCTCTCCCTGTTTCCAGA 121
QY 40 LysGlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 122 AAACAAACCTTGACATTACTTCCATTGCTAGCAACGGCGGAAGTTCATGATGCG 181
QY 60 ValTrpHisMetSerHisProAlaLeuThrGlnLeuArg-----AlaLeuArgTyrCys 77
Db 182 GTGTGG-----CCACCAATTAAACAAGAAAGTAGAGACACTCTCATACCTT 229
QY 78 LysGluIleProAlaLeuAspProGlnLeuLeuAsp----- 89
Db 230 CCCGATTGTAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTACCTGTGAAAAAT 283
QY 90 ---TrpLeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThrVal 108
Db 284 GGATGGGTCTCTGCTGGAATTCAGACTGAGCCT----- 319
QY 109 SerValThrMetIleArgGluGlyPheVal-----GluGlnAsnGluIleProGluGlu 126
Db 320 -----GGATTCTGTACCGTGAACACACACAGCTCACCAGGATAT 358
QY 127 LeuProLeuLeuProLysGluSerArgTyrTrpLeuArg-----GluIleLeuLeu 143
Db 359 TAT-----GATGGCAGATACTGGACCATGTGGAAGTTGCCCATGTTCGGG 403
QY 144 CysAlaAspGlyGluProTrpLeuAla 152
Db 404 TGCACCTGATGCACCTCAGGTGTGGCT 430

RESULT 6
US-11-218-305-17631
; Sequence 17631, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17631
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17631

Alignment Scores:
Pred. No.: 6.77e-18 Length: 583
Score: 238.50 Matches: 85
Percent Similarity: 49.5% Conservative: 20
Best Local Similarity: 40.1% Mismatches: 47
Query Match: 20.6% Indels: 62
DB: 7 Gaps: 11

US-10-718-311-8 (1-227) x US-11-218-305-17631 (1-583)
QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 33 ATGGCTTCTCAGTTATGCTCCTCCGCGCTGTGGCTAC-----TCACCGGCTCAAGCC 86
QY 21 SerMetValAlaProPheThrGlyLeuIleYssSerAlaThrPheProValThrLysLys 40
Db 87 ACCATGGTGTGCTCATTACCGGCTTGAAGTCATCCGCTGCGTATCCCAAGTACCGGCAA 146
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 147 AGCAACACTGATATTAATTCATTGCAACGAGCAACGAGGAGAGATTAGCTGATGAAGGTG 206
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 207 TGGCCACACAGTCG-----GAAAGAAAGAAAGTTTGAGACT-CTCTCTTATCT 250
QY 80 eProAlaLeu---AspProGlnLeu-----LeuAspTrpLeuLeuGluAspSe 96
Db 251 TCCTGACCTTACTGACGTTGAATTGGCCAAAGAAAGTTGACTACCTTCTCCGCAACAAGTG 310
QY 96 rMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGln 115
Db 311 GATTCCCTTGTGTTGAATTTGAG-----TTGGAGCA 340
QY 115 uGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGln 133
Db 341 CGGATTTGTGACCGTGAGCATGGAAGCACACCTGGATAC-----GA 385
QY 133 userArgTyrTrp----- 137
Db 386 TGGCGGTTATTGGCAATGTGGAAGCTTCCTTTGTCGGATGCACTGACTCTGCTCAAGT 445
QY 138 -LeuArgGluIleLeuLeuCysAlaAspGlyGluPro-----Tr 150
Db 446 GTTCAAGGAAGTGAAGAGTGCAA-GACGGAGTAGTACCTTAACGCCCTTTATTAGAATCAT 504
QY 150 pLeuAlaGlyArgThrValValProValSerThrLeuSerGlyProGluLeuAlaLeuGln 170
Db 505 GATTTCGACAAACACCG-----TCAGTCCAGTGCATCAGTTTCA 543
QY 170 nLysLeuGlyLysThrProLeuGlyArgTyrLeu 181
Db 544 TCGCCTACACGCCACCAAGCTTTCACCGGGTGCTT 577

RESULT 7
US-11-247-587-129
; Sequence 129, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 129
; LENGTH: 727
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
```


Qy	79	GlulleProAlaLeuAsp----	ProGlnLeuLeuLeuAspTrpLeuLeuLeuGluAspSerMet	97
Db	267	GACCTTACCGAAGTTGAATTGGTAAAGGAAGACCTTCTCCGCAACAAGTGGATT	326	
Qy	98	---ThrLysAArgPheGluGlnGlnGlyThrValSerValThrMetIleAArgGluGly	116	
Db	327	CCTGTGTTGAATTGCAG-----	TTGGAGACCGGA	356
Qy	117	PheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSer	134	
Db	357	TTTGTGTACCGTGAGCAACGGAACCCCGGATACTAC-----	GATGGC	401
Qy	135	ArgTyrTrp-----	Leu	138
Db	402	CGTTACTGGACAATGTGGAGCTTCCTTGTTCGGATGCACCTGACTCTGCTCAAGTTTG	461	
Qy	139	ArgGluIleLeuLeuCysAlaAspGlyGluPro	149	
Db	462	AAGGAAGTCCAAGAATGCAAAACGGAGTACCCT	494	

RESULT 10
 US-10-953-349-10521
 ; Sequence 10521, Application US/10953349
 ; Publication No. US20060107345A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALEXANDROV, Nickolai et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLY
 ; TITLE OF INVENTION: ENCODED THERBY
 ; FILE REFERENCE: 2750-1579PUS2
 ; CURRENT APPLICATION NUMBER: US/10/953,349
 ; CURRENT FILING DATE: 2004-09-30
 ; NUMBER OF SEQ ID NOS: 40252
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 10521
 ; LENGTH: 718
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-953-349-10521

```

Qy 115 uGlypheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLyseG1 137
Db 340 CGGATTGTGTACCGTGAGCAACGGAAACACTCCCGGATACTAC-----GA 384
Qy 133 userArgTyTrrp----- 137
Db 385 TGGACGGTACTGGACAATGTGGAAGCTTCCATTGTTCCGATGCACGACTCTGCTCAAGT 444
Qy 138 -LeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 445 ATTCAGGAAGATTCAAGAAATGCAAGAAGGAGTACCCGGCGCGCTTCAATAGGATCATC 502

RESULT 11
US-11-218-305-17634
; Sequence 17634, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McQuaid, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; TITLE OF INVENTION: Corn.
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17634
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-17634

```


Db 315 CCTGTGTTGAATTCAG-----TTGGAGCAGCA 344
QY 117 pheVal-----GluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSer 134
Db 345 TTTGTGTACCGTGAGCAGCAAGCAGCCCGGACTACTAC-----GATGGC 389
QY 135 ArgTyrTrp-----Leu 138
Db 390 CGTTACTGACAAATGGGAAGCTTCCTTGTTCGGAGCACTGCTCTCTCAAGTGTTG 449
QY 139 ArgGluLeuLeuCysAlaAspGlyGluPro 149
Db 450 AAGGAAGTCCAAGAATGCCAACTGAGTACCCT 482

RESULT 12

US-11-216-545-4173
; Sequence 4173, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: Mc Laird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; PRIOR FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4173
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Glycine max
US-11-216-545-4173

Alignment Scores:
Pred. No.: 1,07e-13 Length: 841
Score: 203.50 Matches: 64
Percent Similarity: 53.1% Conservative: 13
Best Local Similarity: 44.1% Mismatches: 37
Query Match: 17.6% Indels: 31
DB: 8 Gaps: 10

US-10-718-311-8 (1-227) x US-11-216-545-4173 (1-841)

QY 1 MetAlaSerSerValIleSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 37 ATGGCTTCTCAATGATCTCTCCCGAGCTGTACCACC---GTCAACCGTCCGGTGCC 93
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 94 GCGATGTTGCTCCATTCACCGGCTCAATCCATCGCTGCTTCC---ACGAGGAAG 150
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 151 ACCAACAATGACATTACCTCCATTTGTCAGCAACGGTGAAGAGTACAATGCATGCAGGTG 210
QY 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
Db 211 TGG-----CCACAATGGCAAGAGTTGAGACTCTTTCCTACTTGCCA 258
QY 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuGluAsp 95
Db 259 GACCTCGAT-----GATGCCAATTTGGCAAGAGTGAATACCTCTTAAGGAAGGA 312
QY 96 SerMet-----ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114
Db 313 TGGATTCTTGTGTTGAATTCGAG-----TTGGAG 342
QY 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLys 132

Db 343 CACGGTTTCGTGTACCGTGAGCACACAGGTCACTCCTGTGATCTAT----- 387
QY 133 GluSerArgTyrTrp 137
Db 388 GATGACGCTACTGG 402
RESULT 13
US-10-953-349-9569
; Sequence 9569, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9569
; LENGTH: 799
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-9569
Alignment Scores:
Pred. No.: 1.29e-13 Length: 799
Score: 202.50 Matches: 58
Percent Similarity: 53.1% Conservative: 19
Best Local Similarity: 40.0% Mismatches: 37
Query Match: 17.5% Indels: 31
DB: 6 Gaps: 9

US-10-718-311-8 (1-227) x US-10-953-349-9569 (1-799)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 47 ATGGCTTCTCTATGCTCTCTCCGCTACTATGGTT-----GCCCTCCGGCTCAGGCC 100
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 101 ACTATGGTGCCTCTCTTCAACGGACTTAAGTCTCCGCTTCCAGCCACCCGCAAG 160
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 161 GCTAACACGACATTAATTCATCCATCACAGCAACGGCGGAAGAGTTAACTCATGCAGGTG 220
QY 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
Db 221 TGG-----CCTCCGATTGGAAGAGAGTTGAGACTCTCTTACCTTCTCT 268
QY 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuGluAsp 95
Db 269 GACCTT-----ACCGATTCCGAATTTGGCTAAGGAAGTTGACTACCTTATCCGCAACAG 322
QY 96 SerMet-----ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114
Db 323 TGGATTCTTGTGTTGAATTCGAG-----TTGGAG 352
QY 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLys 132
Db 353 CACGATTGTTGATCGTGACCGTAACTCACCCTGATCTAT----- 397
QY 133 GluSerArgTyrTrp 137
Db 398 GATGACGCTACTGG 412
RESULT 14
US-11-218-305-17632
; Sequence 17632, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:

```
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: Mc Laird, Paul L.
; APPLICANT: Tao, Nengsheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17632
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17632

Alignment Scores:
Pred. No.: 1,368-13 Length: 828
Score: 202.50 Matches: 58
Percent Similarity: 53.1% Conservative: 19
Best Local Similarity: 40.0% Mismatches: 37
Query Match: 17.5% Indels: 31
DB: 7 Gaps: 9

US-10-718-311-8 (1-227) x US-11-218-305-17632 (1-828)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db ATGGCTTCTCTATGCTCTCTTCGCTACTATGGTT-----GCCTCTCCGGCTCAGGCC 103
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db ACTATGGTCGGCTCTTCAACGGACTTAAGTCTCCTCGCTGCTTCCCGAGCCACCGCAAG 163
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db GCTAACACGACGATTACTTCCATCAAGCAACGGCGGAAGATTAACTGATGTCAGGTG 223
Qy 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
Db TGG-----CCTCCGATTGGAAGAAAGATTGTGAGACTCTCTCTTACCTTCT 271
Qy 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuGluAsp 95
Db GACCTT-----ACCGATTCCGAATTGGCTAAGGAAGTTGACTACCTTATCCGCAACAG 325
Qy 96 SerMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114
Db TGGATTCTCTGTGTGAATTTCGAG-----TTGGAG 355
Qy 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuProLys 132
Db CACGATTGTGTACCGTGACGACCGTAACCTACCCCGGATACTAT----- 400
Qy 133 GluSerArgTyrTrp 137
Db GATGGACGCTACTGG 415

RESULT 15
US-11-216-545-4174
; Sequence 4174, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: Mc Laird, Paul L
; APPLICANT: Tao, Nengsheng
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
```

```
; CURRENT APPLICATION NUMBER: US/11/216,545
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4174
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)-(12)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-4174

Alignment Scores:
Pred. No.: 1,858-13 Length: 1023
Score: 202.50 Matches: 64
Percent Similarity: 53.1% Conservative: 13
Best Local Similarity: 44.1% Mismatches: 37
Query Match: 17.5% Indels: 31
DB: 8 Gaps: 10

US-10-718-311-8 (1-227) x US-11-216-545-4174 (1-1023)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db ATGGCTTCTCTCAATGACTCTTCCCGAGCTGTACCACCT---GTCAACCGTCCGGTGCC 284
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db GGATGTGTGCTTCATTCACCTGGCTCAAGTCCATGGCTGCTTCCCC---ACCAGGAAG 341
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db ACCAACCAATGACATTACCTCCATTGCTAGCAACGGTGAAGAGTGCAATGTCATGCAGGTG 401
Qy 61 TrpHisMetSerHisProAlaLeu-----ThrGlnLeuArgAlaLeuArgTyrCysLys 78
Db TGG-----CCACCAGTTGGCAAGAAAGTTTGGAGACTCTTCTTCTTACCTGCCA 449
Qy 79 GluIleProAlaLeuAspProGlnLeu-----LeuAspTrpLeuLeuGluAsp 95
Db GACCTTGAT-----GATGCACAATTGSCAAAGGAAGTAGAATACTTCTTAGGAAGGA 503
Qy 96 SerMet---ThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArg 114
Db TGGATTCTCTGTGGAATTTCGAG-----TTGGAG 533
Qy 115 GluGlyPheVal-----GluGlnAsnGluIleProGluGluLeuProLeuProLys 132
Db CACGGTTCTGTGTCACCGTGAGCACACAGGTTCACCGAGTACTACTAC----- 578
Qy 133 GluSerArgTyrTrp 137
Db GATGGACGCTACTGG 593
```

Search completed: July 11, 2006, 21:05:58
Job time : 137.1 secs

GenCore version 5.1.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 20:23:26 ; Search time 1856.55 Seconds
 (without alignments)
 2525.695 Million cell updates/sec

Title: US-10-718-311-8
 Perfect score: 1159
 Sequence: 1 MASSVISAANVATNSVTAQ.....RLSGKPLLLTELFLPASPLY 227

Scoring table: BLOSUM62
 Xgapop 10.0 , Xgapext 0.5
 Ygapop 10.0 , Ygapext 0.5
 Fgapop 6.0 , Fgapext 7.0
 Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues
 37784340

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Command line parameters:
 -MODEL=frame+ p2n.model -DEV=xlh
 -Q=/abss/ABSSWEB.spool/US10718311/runat_11072006_111531_5371/app.query.fasta_1
 -DB=Published Applications NA Main -QFMT=fastap -SUFFIX=p2n.rnpbm
 -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
 -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss08
 -USER=US10718311@CN 1.1.2326 @runat_11072006_111531_5371 -NCPU=6 -ICPU=3
 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
 -WAE TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1159	100.0	684	3	US-09-855-341-7 Sequence 7, Appli

2	1159	100.0	684	3	US-09-896-8668-16	Sequence 16, Appl
3	1159	100.0	684	7	US-10-359-369-41	Sequence 41, Appl
4	1159	100.0	684	8	US-10-699-050-16	Sequence 16, Appl
5	1159	100.0	684	8	US-10-718-311-7	Sequence 7, Appl
6	1159	100.0	684	9	US-10-462-162-33	Sequence 33, Appl
7	862	74.4	3452	9	US-10-462-162-47	Sequence 47, Appl
8	854	73.7	495	3	US-09-855-341-3	Sequence 3, Appl
9	854	73.7	495	3	US-09-896-8668-12	Sequence 12, Appl
10	854	73.7	495	3	US-10-359-369-37	Sequence 37, Appl
11	854	73.7	495	8	US-10-699-050-12	Sequence 12, Appl
12	854	73.7	495	8	US-10-718-311-3	Sequence 3, Appl
13	854	73.7	498	9	US-10-462-162-29	Sequence 29, Appl
14	851	73.4	1971	9	US-10-462-162-54	Sequence 54, Appl
15	813	70.1	1207	10	US-10-450-763-26139	Sequence 26139, A
16	277.5	23.9	504	10	US-10-487-901-12707	Sequence 7207, Ap
17	276.5	23.9	489	10	US-10-487-901-3503	Sequence 3503, Ap
18	276.5	23.9	718	10	US-10-487-901-5765	Sequence 5765, Ap
19	276.5	23.9	736	10	US-10-487-901-1704	Sequence 1704, Ap
20	276.5	23.9	736	10	US-10-487-901-5043	Sequence 5043, Ap
21	269	23.2	204	3	US-09-839-477-3	Sequence 3, Appl
22	269	23.2	204	8	US-10-758-064-3	Sequence 3, Appl
23	269	23.2	297	10	US-10-487-901-7209	Sequence 7209, Ap
24	269	23.2	377	10	US-10-487-901-3504	Sequence 3504, Ap
25	267.5	23.1	704	9	US-10-425-115-85509	Sequence 85509, A
26	266	23.0	668	10	US-10-487-901-7208	Sequence 7208, Ap
27	253.5	21.9	683	10	US-10-487-901-1713	Sequence 1713, Ap
28	253.5	21.9	737	10	US-10-487-901-5045	Sequence 5045, Ap
29	253.5	21.9	737	10	US-10-487-901-5766	Sequence 5766, Ap
30	247.5	21.4	738	10	US-10-487-901-1716	Sequence 1716, Ap
31	247.5	21.4	753	10	US-10-487-901-5764	Sequence 5764, Ap
32	247.5	21.4	754	10	US-10-487-901-5028	Sequence 5028, Ap
33	246.5	21.3	714	10	US-10-487-901-1703	Sequence 1703, Ap
34	246.5	21.3	736	10	US-10-487-901-5044	Sequence 5044, Ap
35	244	21.1	174	3	US-09-854-286-15	Sequence 15, Appl
36	236.5	20.4	654	10	US-10-487-901-5042	Sequence 5042, Ap
37	235	20.3	169	6	US-10-165-420-3	Sequence 3, Appl
38	230.5	19.9	629	10	US-10-487-901-225	Sequence 225, App
39	228.5	19.7	724	3	US-09-910-664-55	Sequence 55, Appl
40	228.5	19.7	724	8	US-10-333-184-54	Sequence 54, Appl
41	228.5	19.7	726	8	US-10-333-184-202	Sequence 202, App
42	228.5	19.7	764	10	US-10-487-901-4119	Sequence 4119, Ap
43	228.5	19.7	771	3	US-09-770-445-918	Sequence 918, App
44	225.5	19.5	1205	3	US-09-770-445-41	Sequence 41, Appl
45	225	19.4	385	3	US-09-732-627A-1176	Sequence 1176, Ap

ALIGNMENTS

RESULT 1
 US-09-855-341-7
 ; Sequence 7, Application US/09855341
 ; Patent No. US2002002715A1
 ; GENERAL INFORMATION:
 ; APPLICANT: VIITANEN, PAUL V.
 ; APPLICANT: MEYER, KNUT
 ; APPLICANT: VAN DYK, DREW
 ; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
 ; FILE OF INVENTION: IN GREEN PLANTS
 ; FILE REFERENCE: BC1015 US NA
 ; CURRENT APPLICATION NUMBER: US/09/855,341
 ; CURRENT FILING DATE: 2001-05-15
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: MICROSOFT OFFICE 97
 ; SEQ ID NO 7
 ; LENGTH: 684
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
 US-09-855-341-7

Alignment Scores: 6.06e-139 Length: 684
 Pred. No.: 1

Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservativity: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-855-341-7 (1-684)

```
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCATTTCTTCAGAGCTGTGTCCACACGAGCAATGTTACACAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180

Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTACACCCCGGTAAACGCACTGCGCTGCGCTGCTATTGTAAGAGATC 240

Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGCGCTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300

Qy 101 PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db 301 TTTGAACAGCAGGGAACCGGTAAAGCGTACGATGATCCGGAAGGGTGTGTCAGCAG 360

Qy 121 AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAATCCCGAAGAACTGCGCTGCTGCGAAGAGTCTCGTTACTGTTACGTGAA 420

Qy 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160
Db 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTGTC 480

Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGGAGCTGCGCTTACAAAATTTGGGTAAACGCCGTTAGGACGCTAT 540

Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTACATCATCGACATTAACCGGACCTTTATTGAGATAGGCGGTGATGCCGGCTG 600

Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGATTAAGCGTAAACCGCTGTTGCTAACAGAACTGTTT 660

Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGCGTCACCGTTGTAC 681
```

RESULT 2

US-09-896-866B-16
; Sequence 16, Application US/09896866B
; Patent No. US20020151002A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylgucose:Malate Sinapoyltransferase Form Malate Conjugates
; TITLE OF INVENTION: Benzoic Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896.866B
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216, 615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protein
US-09-896-866B-16

Alignment Scores:

Pred. No.: 6.06e-139 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservativity: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-896-866B-16 (1-684)

```
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCATTTCTTCAGAGCTGTGTCCACACGAGCAATGTTACACAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTTCACCTGCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180

Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTACACCCCGGTAAACGCACTGCGCTGCGCTGCTATTGTAAGAGATC 240

Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGCGCTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300

Qy 101 PheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db 301 TTTGAACAGCAGGGAACCGGTAAAGCGTACGATGATCCGGAAGGGTGTGTCAGCAG 360

Qy 121 AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAATCCCGAAGAACTGCGCTGCTGCGAAGAGTCTCGTTACTGTTACGTGAA 420

Qy 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160
Db 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTGTTCTGTGTC 480

Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGGAGCTGCGCTTACAAAATTTGGGTAAACGCCGTTAGGACGCTAT 540

Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTACATCATCGACATTAACCGGACCTTTATTGAGATAGGCGGTGATGCCGGCTG 600

Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGATTAAGCGTAAACCGCTGTTGCTAACAGAACTGTTT 660

Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGCGTCACCGTTGTAC 681
```

RESULT 3

US-10-359-369-41
; Sequence 41, Application US/10359369
; Publication No. US20030215927A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.

; APPLICANT: Viitanen, Paul
; APPLICANT: Meyer, Knut
; APPLICANT: Van Dyk, Drew
; TITLE OF INVENTION: UDP-Glucosyltransferases
; FILE REFERENCE: CL1821 US NA
; CURRENT APPLICATION NUMBER: US/10/359,369
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 60/355,511
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 41
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(684)
US-10-359-369-41

Alignment Scores:
Pred. No.: 6,06e-139 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0

US-10-718-311-8 (1-227) x US-10-359-369-41 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCAATTTCTTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAATCTTCAGCCACATTTCCCTGTTACAAAGAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCGAGTG 180
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
DB 181 TGGCATATGTACACCCCGGTTACGCAACTGCGTGGCTGCGCTATTTGTAAGAGATC 240
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
DB 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300
QY 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
DB 301 TTGTAACAGCAGGAAAAACGGTAAGCGTGACATGATCCGCGAAGGGTTTGTGCGAGCAG 360
QY 121 AsnGluIleProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
DB 361 AATGAATCCCCGAAGACTGCGCGTGCCTGCGCAAGAGCTCTGTTACTGGTTACGTGAA 420
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160
DB 421 ATTTGTTATGTGCCATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTCTCTGTGTCA 480
QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
DB 481 ACGTTAAGCGGCGGAGCTGGCGTTACAAAAATTGGGTAAAAACGCGGTTAGGACGCTAT 540
QY 181 LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
DB 541 CTGTTCAACATCATCGACATTAAACCCGGGACTTTATTGATAGGCGCGTATGCCGGCTG 600
QY 201 TrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
DB 601 TGGGGCGACGTTCCCGCCTCGCATTAAGCGGTAAACCGCTGTTGCTAACAGACTGTTT 660

QY 221 LeuProAlaSerProLeuTyr 227
DB 661 TTACCGGGGTACCCGTTGTAC 681

RESULT 4

US-10-699-050-16
; Sequence 16, Application US/10699050
; Publication No. US20040142437A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate
; TITLE OF INVENTION: Benzoic Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/10/699,050
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei
US-10-699-050-16

Alignment Scores:

Pred. No.: 6,06e-139 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-718-311-8 (1-227) x US-10-699-050-16 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCAATTTCTTTCAGCAGCTGTTCACACAGCAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTCACTGGTCTCAATCTTCAGCCACATTTCCCTGTTACAAAGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAAACCTTGACATCACTTCCATTCCTAGCAATGGTGGAAAGAGTTAGCTGCATGCGAGTG 180
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
DB 181 TGGCATATGTACACCCCGGTTAAGCAACTGCGTGGCTGCGCTATTTGTAAGAGATC 240
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
DB 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300
QY 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
DB 301 TTGTAACAGCAGGAAAAACGGTAAGCGTGACATGATCCGCGAAGGGTTTGTGCGAGCAG 360
QY 121 AsnGluIleProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
DB 361 AATGAATCCCCGAAGACTGCGCGTGCCTGCGCAAGAGCTCTGTTACTGGTTACGTGAA 420
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSer 160

Db 421 ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGGCGGTCTGTACCGTCTGTTCTCTGTGTCA 480
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGCGGATGCGCTTCAAAAAATTTGGGTAAACCGCTTAGGACGCTAT 540
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTTCATCATCATGACATTAAACCGGACCTTATTGAGATAGGCGGTGATGCCGGGCTG 600
Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTGCTTAACAGAACTGTTT 660
Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGCGTCACCGTTGTAC 681

RESULT 5

US-10-718-311-7
; Sequence 7, Application US/10718311
; Publication No. US20040143867A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUIT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/10/718,311
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/855,341
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-10-718-311-7

Alignment Scores:

Pred. No.: 6,06e-139 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-718-311-8 (1-227) x US-10-718-311-7 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValIleThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTGTGCACACGCGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCCTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTGCACACCGCGGTAAACGCAACTGCGCTGCGCTATTTGTAAGAGATC 240
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGCGCTGGATCCGCAACTGCTGCAGCTGGCTGTGTTGCTGGAGGATTCCATGACAAACGT 300

Qy 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleLeuArgGluGlyPheValGluGln 120
Db 301 TTTGAAACGACGAGGAAAAACGTTAAGGTGACGATGATCCCGAAAGGGTTTGTGAGGACAG 360
Qy 121 AsnGluIleProGluGlnLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAA 420
Qy 141 IleLeuLeuCysAlaAspGlyLupProTrpLeuAlaGlyArgThrValValProValSer 160
Db 421 ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGCCTGCTGACCGTCTCTGTGTCA 480
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGGAGCTGCCGTTACAAAAATTTGGGTAAACCGCGCTTAGGACGCTAT 540
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTTCATCATCATGACATTAAACCGGACCTTATTGAGATAGGCGGTGATGCCGGGCTG 600
Qy 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTGCTTAACAGAACTGTTT 660
Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGCGTCACCGTTGTAC 681

RESULT 6

US-10-462-162-33
; Sequence 33, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462,162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
; ORGANISM: Escherichia coli
US-10-462-162-33

Alignment Scores:
Pred. No.: 6,06e-139 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 9 Gaps: 0

US-10-718-311-8 (1-227) x US-10-462-162-33 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValIleThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTGTGCACACGCGCAATGTTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
Qy 41 GlnAenLeuAspIleThrSerIleAlaSerAsnGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCCTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80

Db 181 TGGCATATGTCACACCCCGCTTAACGCAACTCGCTGCGCTATTGTAAAGAGATC 240
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGCCTCGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCCATGACAAACGT 300
Qy 101 PheGluGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db 301 TTGAACACAGCAGGAGAAAACGGTAAGCGTGACGATATCCGGAAGGGTTGTGCGAGCAG 360
Qy 121 AsnGluLeuProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 ATGAATATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGGTTACGTGAA 420
Qy 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160
Db 421 ATTGTGTTATGTCGCGATGTAACCGTGGCTTGCCTGCGTACCGCTGCTGCTGTGTC 480
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTAT 540
Qy 181 LeuPheThrSerThrLeuThrArgAspPheIleGluLeuIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTACATCATCGACATTAACCGGGACTTTATTGAGATAGGCGGTGTCGCGGGCTG 600
Qy 201 TrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT 660
Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGGCGTCACCGCTGTGAC 681

RESULT 7

US-10-462-162-47
; Sequence 47, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462,162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 47
; LENGTH: 3452
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of nucleic acid fragment encoding CPL (from
; OTHER INFORMATION: Escherichia coli), PHBA 1-hydroxylase (from Candida
; OTHER INFORMATION: parapsilosis), and UGT2B1 (from Arabidopsis thaliana).
US-10-462-162-47

Alignment Scores:
Pred. No.: 1 57e-99 Length: 3452
Score: 862.00 Matches: 166
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 74.4% Indels: 0
DB: 9 Gaps: 0

US-10-718-311-8 (1-227) x US-10-462-162-47 (1-3452)

Qy 62 HisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuPro 81
Db 1 CATATGTCACACCCCGCTTAACGCAACTCGCTGCGCTATTGTAAAGAGATCCCT 60

Qy 82 AlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPhe 101
Db 61 GCCTCGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGATTCCATGACAAACGTTTT 120
Qy 102 GluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsn 121
Db 121 GAACAGCAGGAGAAAACGGTAAGCGTGACGATATCCGGAAGGGTTGTGCGAGCAGAA 180
Qy 122 GluLeuProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluLeu 141
Db 181 GAAATCCCGGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGGTTACGTGAAAT 240
Qy 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThr 161
Db 241 TTGTATTATGTCGCGATGTAACCGTGGCTTGCCTGCGTACCGCTGCTGCTGTCAACG 300
Qy 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181
Db 301 TTAAGCGGGCGGAGCTGGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTATCTG 360
Qy 182 PheThrSerSerThrLeuThrArgAspPheIleGluLeuIleGlyArgAspAlaGlyLeuTrp 201
Db 361 TTCACATCATCGACATTAACCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTGTGG 420
Qy 202 GlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeu 221
Db 421 GGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTTA 480
Qy 222 ProAlaSerProLeuTyr 227
Db 481 CCGGGCGTCACCGCTGTGAC 498

RESULT 8

US-09-855-341-3
; Sequence 3, Application US/09855341
; Patent No. US20020002715A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: E. coli
US-09-855-341-3

Alignment Scores:
Pred. No.: 9 77e-100 Length: 495
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-855-341-3 (1-495)

Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82
Db 1 ATGTACACCCCGCTTAACGCAACTCGCTGCGCTATTGTAAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAACCGTTTTTGA 120

```
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
Db 121 CAGCAGGGGAAAACCGTAGCGTAGCATGATCCCGAAGGGTTTGTGACGACGAGATGAA 180
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142
Db 181 ATCCCGGAAGAACTCCGCTGTCGCGAAGAGTCTCGTTACTGCTTACGTTGAAATTTTG 240
Qy 143 LeuProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142
Db 181 ATCCCGGAAGAACTCCGCTGTCGCGAAGAGTCTCGTTACTGCTTACGTTGAAATTTTG 240
Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162
Db 241 TTATGTGCGGATGGTGAACCGGTGCTTACAAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTTC 360
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGCGCGAGCTGCGCTTACAAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTTC 360
Qy 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGACTTTATTTGAGATAGGCGCGTGTAGCGGGCTGTGGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTGTGCTAACAGAACTGTTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTATC 495
RESULT 9
US-09-896-866B-12
; Sequence 12, Application US/0989866B
; Patent No. US20020151002A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Synapoylglucose:Malate Synapoyltransferase Form Malate Conjugates
; FILE REFERENCE: Benzic Acid Glucosides
; CURRENT APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-896-866B-12
Alignment Scores:
Pred. No.: 9,77e-100 Length: 495
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-8 (1-227) x US-09-896-866B-12 (1-495)
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82
Db 1 ATGTACACACCCCGCGTTAAACGCAACTGCGTGGCTGCTATTGTTAAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTGCTCGATGCGTGTGTGGAGGATCCATGACAAAAACGTTTGAA 120
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
Db 121 CAGCAGGGGAAAACCGTAGCGTAGCATGATCCCGAAGGGTTTGTGACGACGAGATGAA 180
```

```
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142
Db 181 ATCCCGGAAGAACTCCGCTGTCGCGAAGAGTCTCGTTACTGCTTACGTTGAAATTTTG 240
Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162
Db 241 TTATGTGCGGATGGTGAACCGGTGCTTACAAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTTC 360
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGCGCGAGCTGCGCTTACAAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTTC 360
Qy 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGACTTTATTTGAGATAGGCGCGTGTAGCGGGCTGTGGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTGTGCTAACAGAACTGTTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTATC 495
RESULT 10
US-10-359-369-37
; Sequence 37, Application US/10359369
; Publication No. US20030215927A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Viitanen, Paul
; APPLICANT: Meyer, Knut
; APPLICANT: Van Dyk, Drew
; TITLE OF INVENTION: UDP-Glucosyltransferases
; FILE REFERENCE: CU1821 US NA
; CURRENT APPLICATION NUMBER: US/10/359,369
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 60/355,511
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(495)
US-10-359-369-37
Alignment Scores:
Pred. No.: 9,77e-100 Length: 495
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 7 Gaps: 0
US-10-718-311-8 (1-227) x US-10-359-369-37 (1-495)
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82
Db 1 ATGTACACACCCCGCGTTAAACGCAACTGCGTGGCTGCTATTGTTAAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTGCTCGATGCGTGTGTGGAGGATCCATGACAAAAACGTTTGAA 120
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
Db 121 CAGCAGGGGAAAACCGTAGCGTAGCATGATCCCGAAGGGTTTGTGACGACGAGATGAA 180
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142
```


Qy	143	LeuCysAlaAspGlyGluProTyrLeuAlaGlyArgThrValValProValSerThrLeu	162
Db	241	TTATGTCCGATGTGAACCGTGGCTTGCCGGTCGTACCGTCTGTGTCAACGTTA	300
Qy	163	SerGlyProGluLeuAlaLeuGlnIysLeuGlyIysThrProLeuGlyArgTyrLeuPhe	182
Db	301	AGCGGGCCGGAGCTGGCGTTACAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTTC	360
Qy	183	ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTyrGly	202
Db	361	ACATCATCGACATTAAACCCGGGACTTTATTGAGATAGCGCGTGATCCGGGCTGTGGGGG	420
Qy	203	ArgArgSerArgLeuArgLeuSerGlyIysProLeuLeuLeuThrGluLeuPheLeuPro	222
Db	421	CGAGTTCCCGCCCTGGCAGTTAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACC	480
Qy	223	AlaSerProLeuTyr	227
Db	481	GCGTCAACCGTTGTAC	495

RESULT 12
US-10-718-311-3
; Sequence 3, Application US/10718311
; Publication No. US20040143867A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUZ
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/10/718,311
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/855,341
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: E. coli
US-10-718-311-3

```

Alignment Scores:
Pred. No.:          9.77e-100      Length:          495
Score:             854.00         Matches:         165
Percent Similarity: 100.0%         Conservative:    0
Best Local Similarity: 100.0%       Mismatches:     0
Query Match:       73.7%           Indels:         0
DB:                8              Gaps:           0

US-10-718-311-8 (1-227) x US-10-718-311-3 (1-495)

```

Qy	63	MesSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla	82
Db	1	ATGTCAACCCCGGTTAAAGCAACTGCGTGCCTATTGTAAAGAGATCCTGCC	60
Qy	83	LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu	102
Db	61	CTGGATCCGCAACTGCCTCGACTGCTGTGTCTGGAGGATTCCATGACAAAACGTTTTGAA	120
Qy	103	GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu	122
Db	121	CAGCAGGGAAAAACCGTAAAGCGTCACGATGATCCCGAAGGCGTTGTTCGAGCAGAAATGAA	180
Qy	123	IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu	142
Db	181	ATCCCCGAAGAACTGCCGCTGCTCCCGAAGAGTCTCGTTACTGGTTACGTGAAATTTTG	240
Qy	143	LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu	162

```
Db 241 TTATGTGCGGATGGTGAACCGTGGCTTGCCGCTGCTACCGTCTTCCCTGTCACGTTA 300
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGGCGGAGCTGGCGCTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTGTT 360
Qy 183 ThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGCCGCTGATGCCGGCTGTGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTAC 495
RESULT 13
US-10-462-162-29
; Sequence 29, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462.162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 29
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-462-162-29
Alignment Scores:
Pred. No.: 9,86e-100 Length: 498
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 9 Gaps: 0
US-10-718-311-8 (1-227) x US-10-462-162-29 (1-498)
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82
Db 1 ATGTGCACACCCCGCTTAACGCACTGGCGCTGCGCTATTGTAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTTGAA 120
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGlnAsnGlu 122
Db 121 CAGCAGGGAAAAACGGTAAGCGTGACGATATCCCGAAGGGTTTGTGAGCAGAAATGAA 180
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142
Db 181 ATCCCCGGAAGAACTGCGCGCTGCGCGAAGAGTCTCGTTACTGTACGTGAATTTTG 240
Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThrLeu 162
Db 241 TTATGTGCGGATGGTGAACCGTGGCTTGCCGCTGCTACCGTCTGTTCTCTGTGTCACGTTA 300
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGGCGGAGCTGGCGCTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTGTT 360
```

```
Qy 183 ThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGCCGCTGATGCCGGCTGTGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTAC 495
RESULT 14
US-10-462-162-54
; Sequence 54, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462.162
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 54
; LENGTH: 1971
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Nucleic acid sequence of a nucleic acid fragment inserted into
; OTHER INFORMATION: expression vector pET19a encoding cPL (from Escherichia coli) and
; OTHER INFORMATION: pHBA 1-hydroxylase (from Cadida parapsilosis).
US-10-462-162-54
Alignment Scores:
Pred. No.: 1,82e-98 Length: 1971
Score: 851.00 Matches: 165
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatches: 1
Query Match: 73.4% Indels: 0
DB: 9 Gaps: 0
US-10-718-311-8 (1-227) x US-10-462-162-54 (1-1971)
Qy 62 HisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIlePro 81
Db 1 CATATGTACACCCCGCTTAACGCACTGGCGTGGCTGCGCTATTTTAAAGAGATCCCT 60
Qy 82 AlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrLysArgPhe 101
Db 61 GCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTT 120
Qy 102 GluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGlnAsn 121
Db 121 GAAACAGCAGGGAAAAACGGTAAGCGTGACGATATCCCGAAGGGTTTGTGAGCAGAAAT 180
Qy 122 GluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIle 141
Db 181 GAAATCCCGAAGAACTGCGCGCTGTCGCGAAGAGTCTCGTTACTGTTTACGTGAAAT 240
Qy 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValProValSerThr 161
Db 241 TTGTTATGTGCGGATGGTGAACCGTGGCTTGCCGCTGTCACGTCGTTCTCTGTGTCACG 300
Qy 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181
Db 301 TTAAGCGGGCGGAGCTGGCGCTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTG 360
Qy 182 PheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrp 201
```

Db 361 TTCACATCATGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTGTGG 420
Qy 202 GlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeu 221
Db 421 GGGCGACGTTCCCGCTCGATTAAAGCGTAAACCGCTGTGCTTAACAGAACTGTTTTTA 480
Qy 222 ProAlaSerProLeuTyr 227
Db 481 CCGGCGTCACCGTTGTAC 498

RESULT 15

US-10-450-763-26139
; Sequence 26139, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 26139
; LENGTH: 1207
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (893)...(925)
; OTHER INFORMATION: 81% homologous to Escherichia coli 4-hydroxybenzoate
; OTHER INFORMATION: synthetase, accession number X57434, Smith-Waterman Score=49.
US-10-450-763-26139

Alignment Scores:

Pred. No.:	7,21e-94	Length:	1207
Score:	813.00	Matches:	164
Percent Similarity:	96.0%	Conservative:	2
Best Local Similarity:	94.8%	Mismatches:	5
Query Match:	70.1%	Indels:	2
DB:	10	Gaps:	0

US-10-718-311-8 (1-227) x US-10-450-763-26139 (1-1207)

Qy 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
Db 634 TGTAAAGAGAGTTCGGCATGTCCACACCCGGGTAAACGCAACTGCGCTGGCTAT 693
Qy 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuLeuAspSer 96
Db 694 TGTAAAGAGATCCCTCGCTCGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
Qy 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db 754 ATGACAAAACGTTTGAACAGCAGGAGGAAAAACGGTAAGCGTGACGATGATCCCGAAGGG 813
Qy 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 814 TTTGTGAGCAGAAATGAAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTAC 873
Qy 137 TrpLeuArgGluLeuLeu-LeuCysAlaAspGlyGluProTrpLeuAla-GlyArgThrV 156
Db 874 TGGTTACGTGAAATTTGTGTATGTGCGCATGGCGAAACCGCGCTTGGCGGGTGTGACCG 933
Qy 156 alValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrP 176
Db 934 TCGTTCTGTGTCAACGTTAAGCGGGCCGGAGCTGGCGTTACAAAAATTTGGGTAAACGC 993

Qy 176 roLeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyA 196
Db 994 CGTTAGGACGCTATCTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCC 1053
Qy 196 iqAspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuL 216
Db 1054 GTGATGCCGGGCTGTGGGGGCGACGTTCCCGCTCGGATTAAAGCGGTAACCGCTGTTC 1113
Qy 216 euThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 1114 TAAACAGAACTGTTTTTACCGGCGTCACCGTTGTAC 1148

Search completed: July 11, 2006, 22:12:04
Job time : 1658.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 20:16:09 ; Search time 200.294 Seconds
(without alignments)
3180.885 Million cell updates/sec

Title: US-10-718-311-8
Perfect score: 1159
Sequence: 1 MASSVISAANATRNVTQA.....RLSGKPLLTFLFPASPLY 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10718311/runat_11072006_111527_5281/app_query.fasta_1
-DB=Issued Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs06p -USER=US10718311@CGN_1_1_381@runat_11072006_111527_5281
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1159	100.0	684	3	US-09-896-866B-16 Sequence 16, Appl
2	1159	100.0	684	3	US-09-855-341-7 Sequence 7, Appli
3	856	73.9	854	3	US-09-064-693A-24 Sequence 24, Appl
4	856	73.9	6641	3	US-09-064-693A-25 Sequence 25, Appl
5	854	73.7	495	3	US-09-896-866B-12 Sequence 12, Appl
6	854	73.7	495	3	US-09-855-341-3 Sequence 3, Appli
7	673	58.1	645	3	US-09-489-039A-1234 Sequence 1234, Ap
8	401	34.6	537	3	US-09-543-681A-3122 Sequence 3122, Ap

Sequence 41, Appl
Sequence 3, Appli
Sequence 8, Appli
Sequence 12, Appl
Sequence 12, Appl
Sequence 8, Appli
Sequence 6, Appli
Sequence 12, Appl
Sequence 6, Appli
Sequence 13, Appl
Sequence 14, Appl
Sequence 11, Appl
Sequence 21, Appl
Sequence 36, Appl
Sequence 34, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 9, Appli
Sequence 5, Appli
Sequence 9, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 5, Appli

US-09-495-797-41
US-09-839-477-3
US-08-152-483B-8
US-08-095-726-12
US-08-096-043-12
US-08-093-577-8
US-08-331-004A-6
US-08-096-623A-12
PCT-US95-13937A-6
US-09-186-002-13
US-09-186-002-14
US-09-441-340-11
US-09-441-340-23
US-09-434-039A-36
US-09-434-039A-34
US-09-182-117-1
US-09-434-039A-1
US-09-182-117-5
US-09-434-039A-5
US-09-182-117-4
US-09-434-039A-4
US-09-441-340-9
US-09-186-002-5
US-08-391-339-9
US-08-484-274A-9
US-09-612-404-9
US-08-090-523-5
US-08-398-627-5
US-08-406-858-5
US-08-476-519-7
US-08-120-703A-5
US-08-399-023-5
PCT-US91-04036-5
PCT-US94-05275-5
PCT-US95-09323-7
US-09-011-151-1
US-09-098-219B-5

ALIGNMENTS

RESULT 1
US-09-896-866B-16
; Sequence 16, Application US/09896866B
; Patent No. 6642036
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapylglucose:Malate Sinapyltransferase Form Malate Conjugate
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896,866B
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei
US-09-896-866B-16

Alignment Scores:
Pred. No.: 3,84e-133 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0


```

; ADDRESSES: Technologies Co.
; STREET: P.O. Box 1625
; CITY: Idaho Falls
; STATE: Idaho
; COUNTRY: USA
; ZIP: 83415-3810
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: Toshiba Satellite Pro T2150CDS
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word Perfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,693A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: W. Gary Goodson
; REGISTRATION NUMBER: 22,387
; REFERENCE/DOCKET NUMBER: LIT-PI-296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (208)526-9469
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-064-693A-24
Alignment Scores:
Pred. No.: 1,31e-95 Length: 854
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-064-693A-24 (1-854)
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrClnLeuArgAlaLeuArgTyr 76
Db 303 TGTAAACGGAGAGTTCGGCATGTGCACACCGCGGTTAACCGCAACTGCGTGGCTGCGCTAT 362
QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 363 TGTAAAGAGATCCCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTC 422
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetLeuArgGluGly 116
Db 423 ATGACAAAACGTTTTGAACAGCAGGAGGAAAAACGGTAAGCGTGACGATGATCCCGGAAGG 482
QY 117 PheValGluClnAsnGluIleProGluLeuLeuProLeuLeuProLysGluSerArgTyr 136
Db 483 TTGTCTGTCAGCAATGAATCCCGGAAAGAACTGCGCGCTGCTGCCGAAAGAGTCTCGTTAC 542
QY 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 543 TGGTTACGTGAATTTTGTATGTGCGGATGTGACCGTGGCTGCGCGTACCGTC 602
QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 603 GTTCTGTGTCAACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTGGGTAAAAACGCG 662
QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg 196
Db 663 TTAGGACGCTATCTGTTTCATCATCATCAATTAAACCGGAGCTTTATTGAGATAGCCGT 722
QY 197 AspAlaGlyLeuTrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
;
Db 723 GATCGCGGCTGTGGGGCGACGTTCCCGCCTGCGATTAAGCGGTAACCGCTGTTGCTA 782
QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 783 ACAGAACTGTTTTTACCGGCGTCCACCGTTGTAC 815

RESULT 4
US-09-064-693A-25
; Sequence 25, Application US/09064693A
; Patent No. 6210937
; GENERAL INFORMATION:
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
; TITLE OF INVENTION: ENGINEERED BACTERIA FOR PRODUCTION
; TITLE OF INVENTION: OF A SPECIFIC PLASTICS PRECURSOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Gary Goodson
; ADDRESSEE: INEL--Lockheed Martin Idaho
; ADDRESSEE: Technologies Co.
; STREET: P.O. Box 1625
; CITY: Idaho Falls
; STATE: Idaho
; COUNTRY: USA
; ZIP: 83415-3810
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: Toshiba Satellite Pro T2150CDS
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word Perfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,693A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: W. Gary Goodson
; REGISTRATION NUMBER: 22,387
; REFERENCE/DOCKET NUMBER: LIT-PI-296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (208)526-9469
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-064-693A-25
Alignment Scores:
Pred. No.: 3,79e-94 Length: 6641
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-064-693A-25 (1-6641)
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrClnLeuArgAlaLeuArgTyr 76
Db 303 TGTAAACGGAGAGTTCGGCATGTGCACACCGCGGTTAACCGCAACTGCGTGGCTGCGCTAT 362
QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 363 TGTAAAGAGATCCCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTC 422
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetLeuArgGluGly 116
Db 423 ATGACAAAACGTTTTGAACAGCAGGAGGAAAAACGGTAAGCGTGACGATGATCCCGGAAGG 482
```

Qy 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 483 TTITGTCGAGCAGAAATGAAATCCCGAAGAACTGCCGCTGCTCCGAAAGAGTCTCGTTAC 542
Qy 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 543 TGGTTACGTGAATTTTGTATGTCGCGATGTTGAACCGTGGCTGGCGGTCGTACCGTC 602
Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 603 GTTCCTGTGTCAACGTTAAGCGGGCGGAGCTGGCGTTACAAAAATTTGGTAAAAACGCCG 662
Qy 177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
Db 663 TTAGAGCGCTATCTGTTACATCATCGACATTAAACCGGAGCTTTATTGAGATAGCGCGT 722
Qy 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
Db 723 GATCGCGGCTGTGGGGCGACGTTCCCGCTGCGATTAAAGCGGTAACCGCTGTTGCTA 782
Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 783 ACAGAACTGTTTTCACCGCGCTCACCGTTGTAC 815
RESULT 5
US-09-896-866B-12
; Sequence 12, Application US/09896066B
; Patent No. 6642036
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugated
; FILE REFERENCE: Benzoic Acid Glucosides
; CURRENT APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-896-866B-12
Alignment Scores:
Pred. No.: 9,49e-96 Length: 495
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-8 (1-227) x US-09-896-866B-12 (1-495)
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82
Db 1 ATGTACACACCCCGCGTTAACGCACTGCGTGGCTGCTATTGTAAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTGA 120
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
Db 121 CAGCAGGGAAAAACCGTAAAGCGTACGATATCCCGAAGGGTTTGTGCGAGCAGAAATGA 180
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluLeu 142
Db 181 ATCCCCGAGAAACTGCGCTGCTCCGAAAGAGTCTCGTTACTGTTACGTGTAATTTTG 240

Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
Db 241 TTATGTGCGGATGGTGAACCGTGGCTTCCCGTGTACCGCTGTTCTCTGTGTCAACGTTA 300
Qy 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGGCGGAGCTGGCGTTACAAAAATTTGGTAAAAACGCCGTTAGGACGCTATCTGTT 360
Qy 183 ThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGAGCTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 420
Qy 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuGluLeuPheLeuPro 222
Db 421 CGAGCTTCCCGCTGCGATTAAAGCGTAAACCGCTGTTGCTAACGAACTGTTTTTACCG 480
Qy 223 AlaSerProLeuTyr 227
Db 481 GCGTCACCGTTGTAC 495
RESULT 6
US-09-855-341-3
; Sequence 3, Application US/09855341
; Patent No. 6683231
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:E. coli
US-09-855-341-3
Alignment Scores:
Pred. No.: 9,49e-96 Length: 495
Score: 854.00 Matches: 165
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 73.7% Indels: 0
DB: 3 Gaps: 0
US-10-718-311-8 (1-227) x US-09-855-341-3 (1-495)
Qy 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82
Db 1 ATGTACACACCCCGCGTTAACGCACTGCGTGGCTGCTATTGTAAAGAGATCCCTGCC 60
Qy 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 61 CTGGATCCGCAACTCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAACGTTTGA 120
Qy 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
Db 121 CAGCAGGGAAAAACCGTAAAGCGTACGATATCCCGAAGGGTTTGTGCGAGCAGAAATGA 180
Qy 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluLeu 142
Db 181 ATCCCCGAGAAACTGCGCTGCTCCGAAAGAGTCTCGTTACTGTTACGTGTAATTTTG 240
Qy 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
Db 241 TTATGTGCGGATGGTGAACCGTGGCTTCCCGTGTACCGCTGTTCTCTGTGTCAACGTTA 300

QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 301 AGCGGCGGAGCTGGCTTACAAAAATTGGGTAAACCGCTTAGGAGCTATCTGTTC 360
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
Db 361 ACATCATCGACATTAAACCGGACCTTTATGAGTAGGCGGTATCCGGGCTGTGGGG 420
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
Db 421 CGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTTACCG 480
QY 223 AlaSerProLeuTyr 227
Db 481 GCGTACCGCTGTGAC 495

RESULT 7

US-09-489-039A-1234
; Sequence 1234, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1234
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1234

Alignment Scores:

Pred. No.: 3,11e-73 Length: 645
Score: 673.00 Matches: 129
Percent Similarity: 87.9% Conservative: 16
Best Local Similarity: 78.2% Mismatches: 20
Query Match: 58.1% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-489-039A-1234 (1-645)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82
Db 148 ATGTCCCATCCTCGCTTACGCGACTGCGTGGCTGGCTATTTTCGGTATGCGCATCC 207
QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
Db 208 CTGCGCGCGCGCTTAAGTACTGCTGCTGAGGACTCAATGACCCGAGCGCTTGAA 267
QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
Db 268 CAACAAGGAAGCAGGTACCGTGGCTGGTTAACGAGGGTATATCGCGCGCTGACGCG 327
QY 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142
Db 328 CTACTGACGACGCGGCGCTGCTGCCCGACGACCGCGCTACTGCTGGCGGAGATTATA 387
QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
Db 388 CTCAATGCCGATGCGGAGCCCTGGCTGGCGCGCGCACGGTGGCGCGGAGTCGACGCTG 447
QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
Db 448 TGTGGTCCGAGCTGGCGCTACAGCAGCTCGGCAAACTCCGCTGGCGCGGATACCTGTTT 507
QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202

Db 508 ACCTCGTCGACATTAAACCGCGATTTATTGAAATTTGTCGCGATGACGCGCTGTGGGG 567
QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
Db 568 CGTGTTCCTCCGCTCGGCTGAGCGCAAAACCCCTGCTGACCGAGCTTTTTTTCCT 627
QY 223 AlaSerProLeuTyr 227
Db 628 GCGTCCCGCTGTGAC 642

RESULT 8

US-09-543-681A-3122
; Sequence 3122, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3122
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3122

Alignment Scores:
Pred. No.: 8.23e-40 Length: 537
Score: 401.00 Matches: 81
Percent Similarity: 68.6% Conservative: 15
Best Local Similarity: 57.9% Mismatches: 44
Query Match: 34.6% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-543-681A-3122 (1-537)

QY 88 LeuAspTrpLeuLeuLeuGluAspSerMetThrLysArgPheGluGlnGlnGlyLysThr 107
Db 109 TTAAGTTGGCTGCTGGAATTAGTTCATGACACGACGCTTTTGAGCAACATTCCTCA 168
QY 108 ValSerValThrMetIleArgGluGlyPheValGluGlnAsnGluLeuProGluGluLeu 127
Db 169 GTACCGGTAATGCCCATATCAAGAAGGTTTTTATTGAATATATCGAGCCTCTGATGA 228
QY 128 ProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeuLeuCysAlaAspGly 147
Db 229 AAGTGTTTACCTTATAGCGACGCTATTTGGCTAAGAGAATCTCTTTTGGGGATAAT 288
QY 148 GluProTrpLeuAlaGlyArgThrValValProValSerThrLeuSerGlyProGluLeu 167
Db 289 GTACCTTGGTTATTAGGCGCAACATTAGTCCGCGAAGACACATTAAACGGGTGAAGATCG 348
QY 168 AlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPheThrSerSerThrLeu 187
Db 349 CAATTAGTGAATTTTGGCTACGCGTCCATTAGGACGTTATCTGTTCAGGAAACAACCTTA 408
QY 188 ThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGlyArgArgSerArgLeu 207
Db 409 AGTCGTGATTATTTATTCATATTGGGCAACAAATGGACATTGTTAGCTGTTCCCGTTTC 468
QY 208 ArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuProLysSerProLeuTyr 227
Db 469 CAGCTTTCAGATAAACCTTTATTATTAACTGAGGTGTTTTTGCCTGCATCACCTGTATAT 528

RESULT 9

US-09-495-797-41
; Sequence 41, Application US/09495797
; Patent No. 6369296
; GENERAL INFORMATION:

```
; APPLICANT: Ratcliff, Frank G
; APPLICANT: Martin-Hernandez, Ana M
; APPLICANT: Baulcombe, David C
; TITLE OF INVENTION: Viral Vectors
; FILE REFERENCE: Mewburn 43,047
; CURRENT APPLICATION NUMBER: US/09/495,797
; CURRENT FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N. benthamiana rubisco partial cDNA sequence
US-09-495-797-41

Alignment Scores:
Pred. No.: 1,25e-24 Length: 499
Score: 277.50 Matches: 71
Percent Similarity: 61.5% Conservative: 17
Best Local Similarity: 49.7% Mismatches: 30
Query Match: 23.9% Indels: 25
DB: 3 Gaps: 7

US-10-718-311-8 (1-227) x US-09-495-797-41 (1-499)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 8 ATGGCTTCCTCAGTTCTTTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 67
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 68 AACATGGTTGCACCTTCACTGGCCTTAAGTCAGCTGCTCATTCCTGTTTCAAGGAG 127
QY 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
Db 128 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCATGCAGGTG 187
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArg-----AlaLeuArgTyrCysLys 78
Db 188 TGG-----CCCAATTAAACAGAAAGATGACGACTCTCTCATACCTTCTCT 235
QY 79 GluIleProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSer 96
Db 236 GATTTGAGC-----CAGGAGCAATTGCTTAGTGAAGTTGAGTACCTTTTGAAAAATGGA 289
QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db 290 TGGGTTCTCTTGTGTAATTCGAG-----ACTGAGCAGCGGA 325
QY 117 PheVal-----GluGlnAenGluIleProGluGluLeuProLeuLeuProLysGluSer 134
Db 326 TTGTCTACCGTGAACCAACCAAGTCAACGAGTACTAT-----GATGGC 370
QY 135 ArgTyrTrp 137
Db 371 AGATACTGG 379

RESULT 10
US-09-839-477-3
; Sequence 3, Application US/09839477
; Patent No. 6723895
; GENERAL INFORMATION:
; APPLICANT: Shortosh, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; FILE REFERENCE: COA-CARBOXYLASE
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)
US-09-839-477-3

Alignment Scores:
Pred. No.: 3.23e-24 Length: 204
Score: 269.00 Matches: 53
Percent Similarity: 96.7% Conservative: 6
Best Local Similarity: 86.9% Mismatches: 2
Query Match: 23.2% Indels: 0
DB: 3 Gaps: 0

US-10-718-311-8 (1-227) x US-09-839-477-3 (1-204)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAenValThrGlnAla 20
Db 1 ATGGCTTCCTCAGTTCTTTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTTGCACCTTCACTGGCCTTAAGTCAGCTGCTCATTCCTGTTTCAAGGAG 120
QY 41 GlnAenLeuAspIleThrSerIleAlaSerAenGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCACTTCCATTGCCAGCAACGGCGGAAGAGTGCATGCAGGTG 180
QY 61 Trp 61
Db 181 TGG 183

RESULT 11
US-08-152-483B-8
; Sequence 8, Application US/08152483B
; Patent No. 5529909
; GENERAL INFORMATION:
; APPLICANT: Della-Cioppa, Guy
; APPLICANT: Kumagai, Monto
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
; TITLE OF INVENTION: PROTEIN FUSION ENZYME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,483B
; FILING DATE: No. 5529909ember 12, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7/857,602
; FILING DATE: March 30, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 923,692
; FILING DATE: July 31, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 600,244
; FILING DATE: October 22, 1990
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 641,617
FILING DATE: January 16, 1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 737,899
FILING DATE: July 26, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: BIOG-20240/8129-040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-3660
TELEFAX: (415) 854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1442
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
DESCRIPTION: NO
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM:
IMMEDIATE SOURCE:
CLONE:
FEATURE:
US-08-152-483B-8

Alignment Scores:
Pred. No.: 5,05e-22 Length: 1442
Score: 262.50 Matches: 56
Percent Similarity: 85.1% Conservative: 7
Best Local Similarity: 75.7% Mismatches: 4
Query Match: 22.6% Indels: 7
DB: Gaps: 1

US-10-718-311-8 (1-227) x US-08-152-483B-8 (1-1442)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 9 ATGGCTTCTCAGTTCTTCTCTGAGCAGTTGCGCCGAGCAATGTTGCTCAAGCT 68
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 69 AACATGGTTGCACCTTCTACTGGCTTAAGTCAGCTGCCTCATTCCTCTTTCAAGGAG 128
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 129 CAAACCTTGACATCACTTCCATTGCGCAACCGGGGAGAGTGAATGCAATG 182
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeu 74
DB 183 -----CCGGAACCTACCCGCTGTCGCGGCTC 209

RESULT 12
US-08-095-726-12
Sequence 12, Application US/08095726
Patent No. 5530188
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B
TITLE OF INVENTION: Beta-Carotene Biosynthesis in
TITLE OF INVENTION: Genetically Engineered Hosts
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amoco Corp., Patents and Licensing Dept

STREET: 200 E Randolph St
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60680-0703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,726
FILING DATE: 21-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/785,566
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5530188val B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3128567180
TELEFAX: 3128564972
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 177 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-095-726-12

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: Gaps: 0

US-10-718-311-8 (1-227) x US-08-095-726-12 (1-177)

QY 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCTCAGTTCTTCTCTGAGCAGTTGCGCCGAGCAATGTTGCTCAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AACATGGTTGCACCTTCTACTGGCTTAAGTCAGCTGCCTCATTCCTCTTTCAAGGAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
DB 121 CAAACCTTGACATCACTTCCATTGCGCAACCGGGGAGAGTGAATGCAATGCAATG 177

RESULT 13
US-08-096-043-12
Sequence 12, Application US/08096043
Patent No. 5530189
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B
TITLE OF INVENTION: Lycopene Biosynthesis in
TITLE OF INVENTION: Genetically Engineered Hosts
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amoco Corp., Patents and Licensing Dept
STREET: 200 E Randolph St
CITY: Chicago
STATE: IL
COUNTRY: USA

```
;
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,043
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/785,568
; FILING DATE: 30-OCT-1991
; TELECOMMUNICATION INFORMATION:
; NAME: Galloway, No. 5530189val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-096-043-12

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x US-08-096-043-12 (1-177)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTCTTCTCTGCGACAGTTGCCACCCGCGAGCAATGTTGCTCAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCGCCTTTCACGTGGCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCATTCCATTGCCAGCAACGCGGAGAGTGCAATGCATGCAG 177

RESULT 14
US-08-093-577-8
; Sequence 8, Application US/08093577
; Patent No. 5545816
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Hwei-Che B
; TITLE OF INVENTION: Phytoene Biosynthesis in
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,577
; FILING DATE: 19-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,569
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5545816val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-093-577-8

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x US-08-093-577-8 (1-177)

Qy 1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCAGTCTTCTCTGCGACAGTTGCCACCCGCGAGCAATGTTGCTCAAGCT 60

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCGCCTTTCACGTGGCCTTAAGTCAGCTGCCTCATTCCTGTTTCAAGGAAG 120

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCATTCCATTGCCAGCAACGCGGAGAGTGCAATGCATGCAG 177

RESULT 15
US-08-331-004A-6
; Sequence 6, Application US/08331004A
; Patent No. 5618988
; GENERAL INFORMATION:
; APPLICANT: Hauptmann, Randal
; APPLICANT: Eschenfeldt, William H
; APPLICANT: English, Jami
; APPLICANT: Brinkhaus, Friedhelm L
; TITLE OF INVENTION: Enhanced Carotenoid Accumulation
; TITLE OF INVENTION: in Storage Organs of Genetically
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corporation, Law Dept
; STREET: 55 Shuman Boulevard, Suite 600
; CITY: Naperville
; STATE: IL
; COUNTRY: USA
; ZIP: 60563-8437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,004A
; FILING DATE:
```

```

; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5618988val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7087172447
; TELEFAX: 7087172430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-331-004A-6

```

```

Alignment Scores:
Pred. No.: 1.81e-22 Length: 177
Score: 254.00 Matches: 51
Percent Similarity: 96.6% Conservative: 6
Best Local Similarity: 86.4% Mismatches: 2
Query Match: 21.9% Indels: 0
DB: 2 Gaps: 0

```

US-10-718-311-8 (1-227) x US-08-331-004A-6 (1-177)

```

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCAGTTCTTTCTCTGCAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AACATGGTGGCGCTTTCACCTGACCTTAAGTCAGCTGCCTCATTCCTCTGTTTCAAGGAAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGln 59
Db 121 CAAACCTTGACATCACTTCCTATTGCCAGACGCGGGAAGAGTGCAATGCATGCAG 177

```

Search completed: July 11, 2006, 21:02:56
Job time : 204.294 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 19:48:48 ; Search time 5789.68 Seconds
 (without alignments)
 3288.703 Million cell updates/sec

Title: US-10-718-311-8
 Perfect score: 1159
 Sequence: 1 MASSVISAANVATRSNVTTQA.....RLSGKPLLLTELFLPASPLY 227

Scoring table: BLOSUM62
 Xgapop 10.0, Xgapext 0.5
 Ygapop 10.0, Ygapext 0.5
 Fgapop 6.0, Fgapext 7.0
 Delop 6.0, Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Command line parameters:
 -MODEL=frame+ p2n.model -DEV=xlp
 -Q=/abs/ABSSWEB spool/US10718311/runat 11072006 111523 5213/app query.fasta_1
 -B=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
 -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p
 -USER=US10718311@CGN_1_12067 @runat 11072006 111523 5213 -NCPU=6 -ICPU=3
 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1:*
 2: gb_est3:*
 3: gb_est4:*
 4: gb_est5:*
 5: gb_est6:*
 6: gb_hic:*
 7: gb_est2:*
 8: gb_est7:*
 9: gb_est8:*
 10: gb_est9:*
 11: gb_gse1:*
 12: gb_gse2:*
 13: gb_gse3:*
 14: gb_gse4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	618	53.3	832	13	CL666350 PRI0152b
2	510	44.0	828	13	CL687382 PRI0146c
3	320	27.6	504	7	AW040748 EST283612
4	316	27.3	431	7	AW623440 EST321385

5	316	27.3	554	1	AI777142	AI777142 EST258107
6	314	27.1	439	7	AW041050	AW041050 EST23914
7	314	27.1	456	7	AW091579	AW091579 EST284855
8	314	27.1	612	7	AW622820	AW622820 EST306806
9	313.5	27.0	420	1	AW041064	AW041064 EST283928
10	313.5	27.0	422	1	AI775340	AI775340 EST256440
11	313	27.0	449	7	AW093612	AW093612 EST286792
12	313	27.0	533	7	AW622747	AW622747 EST306884
13	313	27.0	621	2	BG129744	BG129744 EST475390
14	312.5	27.0	416	7	AW094716	AW094716 EST287896
15	312	26.9	434	3	BP896321	BP896321 BP896321
16	312	26.9	436	1	AI773402	AI773402 EST254502
17	312	26.9	436	7	AW039480	AW039480 EST281761
18	312	26.9	437	7	AW096716	AW096716 EST289896
19	312	26.9	439	7	AW039685	AW039685 EST282133
20	312	26.9	441	7	AW220510	AW220510 EST297063
21	312	26.9	443	7	AW037399	AW037399 EST278901
22	312	26.9	446	1	AI773933	AI773933 EST255033
23	312	26.9	446	1	AI899617	AI899617 EST269060
24	312	26.9	446	7	AW038809	AW038809 EST280765
25	312	26.9	446	7	AW039508	AW039508 EST281789
26	312	26.9	446	7	AW093842	AW093842 EST287022
27	312	26.9	447	2	BI929385	BI929385 EST549274
28	312	26.9	447	7	AW615833	AW615833 EST325283
29	312	26.9	449	1	AI772486	AI772486 EST253586
30	312	26.9	449	7	AW039228	AW039228 EST281463
31	312	26.9	449	7	AW094548	AW094548 EST287728
32	312	26.9	449	7	AW399384	AW399384 EST309884
33	312	26.9	452	7	AW037501	AW037501 EST276722
34	312	26.9	452	7	AW039295	AW039295 EST281522
35	312	26.9	453	7	BE462707	BE462707 EST325036
36	312	26.9	454	1	AI484250	AI484250 EST248477
37	312	26.9	454	1	AI490880	AI490880 EST241589
38	312	26.9	454	7	AW037928	AW037928 EST279572
39	312	26.9	456	1	AI772090	AI772090 EST253190
40	312	26.9	456	1	AI779629	AI779629 EST260508
41	312	26.9	470	3	BP889601	BP889601 BP889601
42	312	26.9	475	3	BP898763	BP898763 BP898763
43	312	26.9	476	3	BP905268	BP905268 BP905268
44	312	26.9	476	3	BP910180	BP910180 BP910180
45	312	26.9	477	3	BP909810	BP909810 BP909810

ALIGNMENTS

RESULT 1

CL666350/c

LOCUS

DEFINITION

CL666350 832 bp DNA linear GSS 09-JUL-2004
 PRI0152b_A03 - PRI0152b_B21 (832) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CL666350.1 GI:50159025

GSS.

Pristionchus pacificus

Pristionchus pacificus

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

Neodiplogasteridae; Pristionchus.

1 (bases 1 to 832)

Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.

AppDB: an AcedB database for the nematode satellite organism

Pristionchus pacificus

Nucleic Acids Res. 32 (1), D421-D422 (2004)

14681447

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spananstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 1. 832
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

FEATURES

source
 location/Qualifiers
 1. 828
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
 Seq primer: T7
 Class: fosmid ends.

source
 location/Qualifiers
 1. 828
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Alignment Scores:
 Pred. No.: 5,678-57 Length: 832
 Score: 618.00 Matches: 120
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 53.3% Indels: 0
 DB: 13 Gaps: 0

US-10-718-311-8 (1-227) x CL666350 (1-832)

QY 108 ValSerValThrMetIleArgGluGlyPheValGluGlnAenGluIleProGluGluLeu 127
 DB 832 GTAACGGTGACGATCATCCGGAAGGGTTTGTGAGCAGAGATGAATCCCGAAGACTG 773
 QY 128 ProLeuLeuProLysGluSerArgTyrTrpLeuArgGluIleLeuLeuCysAlaSpGly 147
 DB 772 CGCGTGCTGCCGAAGAGTCTCGTTACTGGTTACGTGAATTTTGTATGTGCCGATGGT 713
 QY 148 GluProTtpLeuAlaGlyArgThrValProValSerThrLeuSerGlyProGluLeu 167
 DB 712 GNACCGTGGCTGCCGGTGTACCGTCTGTCTGTCAAGTTAAGCGGGCGGAGCTG 653
 QY 168 AlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPheThrSerSerThrLeu 187
 DB 652 GCGTTACAAAATTTGGTAAACCGCGTTAGGACGCTATCTGTTACATCATCGACATTA 593
 QY 198 ThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTtpGlyArgSerArgLeu 207
 DB 592 ACCCGGAGCTTTATTAGATAGCGCGTGATGCCGGGCTGTGGGGCGACGTTCCCGCGCTG 533
 QY 208 ArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuProAlaSerProLeuTyr 227
 DB 532 CGATTAAAGCGGTAAACCGCTGTGTCTACAGAACTGTTTACCGGCGTCAACCGTTGTAC 473

RESULT 2

LOCUS CL687382 828 bp DNA linear GSS 09-JUL-2004
 DEFINITION PRI0146c_H03.2 - PRI0146c_BR (828) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION CL687382

VERSION CL687382.1 GI:50196135

KEYWORDS GSS.

SOURCE Pristionchus pacificus

ORGANISM Pristionchus pacificus

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

Neodiplogasteridae; Pristionchus.

REFERENCE 1 (bases 1 to 828)

AppaDB: an AcedB database for the nematode satellite organism

AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.

TITLE Pristionchus pacificus

JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)

PUBMED 14681447

COMMENT Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: raif.sommer@tuebingen.mpg.de

ORIGIN

Alignment Scores:
 Pred. No.: 4,26e-45 Length: 828
 Score: 510.00 Matches: 99
 Percent Similarity: 95.2% Conservative: 1
 Best Local Similarity: 94.3% Mismatches: 5
 Query Match: 44.0% Indels: 0
 DB: 13 Gaps: 0

US-10-718-311-8 (1-227) x CL687382 (1-828)

QY 57 CysMetGlnValTtpHisMetSerHisProAlaLeuThrGlnLeuAlaLeuArgTyr 76
 DB 514 TGTAAACGGAGAGTTCGGCATGTCAACCCCGGTTAAGCAACTGCGTGGCGCTAT 573
 QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTtpTrpLeuLeuGluAspSer 96
 DB 574 TTTAAAGAGATCCCTGCCCTGGATCCGCAACTGTCGACTGGCTGTTCTGGAGGATTCC 633
 QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
 DB 634 ATGCAAAACGTTTGAACAGCAGGGAACCGTAACGCGTACGATATCCGCAAGGG 693
 QY 117 PheValGluGlnAenGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
 DB 594 TTTGTGAGCAGAGATGAATCCCGAGAACTGCGGCTGCTGCCGAAGAGTCTGTTAC 753
 QY 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTtpTrpLeuAlaGlyArgThrVal 156
 DB 754 TGGTTACGTGAATTTTGTATGTGCCATGTTGAACCGTGAACCGTGGCTGCGCGTACCGCTC 813
 QY 157 ValProValSerThr 161
 DB 814 GTTCTGTGTCAACG 828

RESULT 3

LOCUS AW040748

DEFINITION EST283612 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

clone cLET8G6, mRNA sequence.

ACCESSION AW040748

VERSION AW040748.1

KEYWORDS EST.

SOURCE Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 504)

AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,

Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,

Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and

Giovannoni,J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

COMMENT Contact: CUGI

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source

REFERENCE
AUTHORS

1 (bases 1 to 431)

AUTHORS	TITLE	JOURNAL	COMMENT.
...

van der Hoeven, R. S., Bezzerides, J. L., Matern, A. L., Holt, I. E.,
Liang, F., Hansen, T., Craven, M. B., Bowman, C. L., Ronning, C. M.,
Nierman, W., Fraser, C. M., Martin, G. B., Giovannoni, J. J. and
Tanksey, S. D.
Generation of ESTs from tomato flower tissue, 3-8 mm buds
Unpublished (1999)
Contact: CUGI

FEATURES

```

3 prime sequence.
Location/Qualifiers
1. 431
/mol_type="mRNA"
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"

```

ORIGIN

Alignment Scores:

Pred. No.:	1.59e-24	Length:	508
Score:	320.00	Matches:	84
Percent Similarity:	60.4%	Conservative:	12
Best Local Similarity:	52.8%	Mismatches:	39
Query Match:	27.6%	Indels:	24
DB:	7	Gaps:	7

US-10-718-311-8 (1-227) x AW040748 (1-504)

1	MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAenValThrGlnAla	20
QY		
12	ATGGCTTCCTCTGTGCTATTCTTCAGCACTGTTGCCACGACGACAAATGTTACAAAGCT	71
Db		
21	SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys	40
QY		
72	AGCATGGTTGCACCTTTCACCTGGTCTCAATCTTTTCAGCACTTTCCTGTTACAAAGAAG	131
Db		
41	GlnAenLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal	60
QY		
132	CAAAACCTTGACATCACCTTCATTGCTAGCAATGGTGGAAAGATGACTCGATCGCAGGTTG	191
Db		
61	TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle	80
QY		
192	TGGCCACCAATTAACTATGAGAGAGTACGAGACACTCTCATACCTTCCTGATTG	245
Db		
81	ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThr	98
QY		
246	TCTTGACGACAAATGCTTAGTGAAATGAGTACCTTTTGAAATAATGATGGGTT	299
Db		
99	LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal	118
QY		
300	CCTTGGTGGAAATTGAGCTGAGACCGGATTTGTC	335
Db		
119	GlulnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr	136
QY		
336	TACCGTGAGAACCAACAGTCACCGAGTACTATGATGGAAGGTAC	380
Db		
137	TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla	152
QY		
381	TGACCATGTGGAAATGCTTGTGGTGGCTGATGCAACCCAGCTTTGGCT	437
Db		

RESULT 4				
AW623440	AW623440	431 bp	mRNA	linear EST i8-MAY-2001
LOCUS	EST321385	tomato flower buds 3-8 mm,	Cornell University	
DEFINITION	Lycopersicon esculentum cDNA clone tCOTB10t24 5', mRNA sequence.			
ACCESSION	AW623440			
VERSION	AW623440.1	GI:7336467		
KEYWORDS	EST.			
SOURCE	Lycopersicon esculentum (Solanum lycopersicum)			
ORGANISM	Lycopersicon esculentum			

ORIGIN

Alignment Scores:

Pred. No.:	3 47e-24	Length:	431
Score:	316.00	Matches:	84
Percent Similarity:	60.4%	Conservative:	12
Best Local Similarity:	52.8%	Mismatches:	39
Query Match:	27.3%	Indels:	24
DB:	7	Gaps:	7

US-10-718-311-8 (1-227) x AW623440 (1-431)

Qy		20	
1	MetaLaseSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla	20	
Db		65	
6	ATGGCTTCCTGTGTCATTCTTTCAGCAGCTGTGCCACGCAGCAATGTTTACA	70	
Qy		40	
21	SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys	40	
Db		125	
66	AGCATGGTTGCACCTTTTCATGGTGCTCAAAATCTTCAGCCACTTTCCTCGTTTACNAAGA	125	
Qy		60	
41	GinAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal	60	
Db		185	
126	CAAACCTTGACATCACCTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG	185	
Qy		80	
61	TrpHisMetSerHisProAlaLeuThrGlnLeuAargAlaLeuArgTyfCysLysGluIle	80	
Db		239	
186	TGGCCACCATAATTAAAC-----ATGAAGAGATPACGAGACACTCTCATACCTTCCTGATTG	239	
Qy		98	
81	ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr	98	
Db		293	
240	TCT-----GAGCAGCAATTGCTTAGTGAAAATTGAGTACCTTTTGAAAAATGGAATGGGTT	293	
Qy		118	
99	LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleAArgGluGlyPheVal	118	
Db		329	
294	CCTTGCTGGAAATTGAG-----ACTGAGCAGCGATTGTCTC	329	
Qy		136	
119	-----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyf	136	
Db		374	
330	TACCGTGAGAACCAACAAGATCCACGAGTACTAT-----GATGGAAGGTAC	374	
Qy		152	
137	TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla	152	

```

Db      375 TGGACCATGTGGAGTTCCTATGTTAGGTGTCAGTGAACCAAGTGTGGCT 431
RESULT 5
AI777142
LOCUS      554 bp      mRNA      linear      EST 18-MAY-2001
DEFINITION EST258107 tomato resistant, Cornell Lycopersicon esculentum cDNA
AI777142
ACCESSION AI777142
VERSION   AI777142.1 GI:5275183
KEYWORDS  EST.
SOURCE    Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM  Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 554)
AUTHORS   D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
TITLE     Generation of ESTs from Pseudomonas resistant tomato
JOURNAL   Unpublished (1999)
COMMENT   Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES             Location/Qualifiers
     source           1..554
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="R11-12 (3S::Pto in Rio Grande x Money Maker)"
                     /db_xref="taxon:4081"
                     /clone="CLER20020"
                     /tissue_type="leaf"
                     /dev_stage="4-week old"
                     /lab_host="SOLR"
                     /clone_lib="tomato resistant, Cornell"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Alignment Scores:
Pred. No.:      5,07e-24      Length:      554
Score:          316.00      Matches:      84
Percent Similarity: 60.4%      Conservative: 12
Best Local Similarity: 52.8%      Mismatches:   39
Query Match:    27.3%      Indels:       24
DB:             1          Gaps:         7

US-10-718-311-8 (1-227) x AI777142 (1-554)

Qy      1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||
Db      1 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTTGCCACACGACGAAATGTTACACAAGCT 60

Qy      21 SerMetValAlaProPheThrGlyLeuIysSerSerAlaThrPheProValThrIysLys 40
|||||
Db      61 AGCATGGTTGGACCTTTCCACCTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
|||||
Qy      41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
|||||
Db      121 CAAACCTTGATCATCTCCATTCCTAGCATGCTGGGAAGAGTTAGCTGATGACGAGTG 180
|||||
Qy      61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
|||||
Db      181 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCCTCGATTG 234
|||||
Qy      81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
|||||

```

```

Db      235 TCT-----GACGAGCAATTCCTTAGTGAATAGTACCTTTTGAAAAAATGGATGGGTT 288
Qy      99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
|||||
Db      289 CCTTGCTTGGAAATTTAG-----ACTGACACGAGATTGTC 324
Qy      119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
|||||
Db      325 TACCGTGAGAACAAAGTCCACAGGATCTAT-----GATGGAAGGTAC 369
Qy      137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
|||||
Db      370 TGGACCATGTGGAAAGTTCCTATGTTGGGTGCATGTGCAACCAAGTGTGGCT 426
RESULT 6
AW041050
LOCUS      439 bp      mRNA      linear      EST 18-MAY-2001
DEFINITION EST283914 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLET11E11, mRNA sequence.
AW041050
ACCESSION AW041050
VERSION   AW041050.1 GI:5899804
KEYWORDS  EST.
SOURCE    Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM  Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 439)
AUTHORS   D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
TITLE     Generation of ESTs from tomato leaf tissue
JOURNAL   Unpublished (1999)
COMMENT   Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES             Location/Qualifiers
     source           1..439
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="Rio Grande PtoR"
                     /db_xref="taxon:4081"
                     /clone="cLET11E11"
                     /tissue_type="leaf"
                     /dev_stage="4-6 week old plants"
                     /lab_host="X11-Blue MRF"
                     /clone_lib="tomato mixed elicitor, BTI"
                     /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisocotinic
acid, BTH, jasmonic acid, ethylene, fenchion, EIX,
okadaic acid, or systemin prior to tissue harvest.
Site was destroyed during cloning."

ORIGIN
Alignment Scores:
Pred. No.:      5.92e-24      Length:      439
Score:          314.00      Matches:      83
Percent Similarity: 59.7%      Conservative: 12
Best Local Similarity: 52.2%      Mismatches:   40
Query Match:    27.1%      Indels:       24
DB:             7          Gaps:         7

US-10-718-311-8 (1-227) x AW041050 (1-439)

Qy      1 MetAlaSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
|||||
Db      9 ATGGCTTCCTGTCATTTCTTCAGCAGCTGTTGCCACACGACGAAATGTTACACAAGCT 68

```

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 |||||
 Db 69 AGCATGGTGGACCTTTCACCTGTCTCAAAATCTTCAGCCACTTTCCTGTGTACAAAGAAG 128
 |||||
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 129 CAAACCTTGACATCACTCCATCTGTAGCAATGGTGGAGAGTTAGTCGATGCAAGTG 188
 |||||
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
 |||||
 Db 189 TGGCCACCAATTAAC-----ATGAAGAGTACGAGACACTCTCATACCTTCCTGATTGTG 242
 |||||
 QY 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
 |||||
 Db 243 TCT-----GACGAGCAATTCCTAGTGAATAGTACCTTTTGAAAAATGATGGGTT 296
 |||||
 QY 99 LysArgPheGlnGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118
 |||||
 Db 297 CCTGTGTTGAATTCAG-----ACTGAGCAGGATTGTG 332
 |||||
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuProLysGluSerArgTyr 136
 |||||
 Db 333 TACCGTGAGAACCAACCAAGTCAACCAAGATACAT-----GATGGAAGGTAC 377
 |||||
 QY 137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
 |||||
 Db 378 TGGACCATGTGAAGTTGGCTATGTTGGTGTCATGTATGATCAACCAAGTGTGGCT 434
 |||||

RESULT 7

LOCUS AW091579
 DEFINITION EST284855 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET15H6, mRNA sequence.

ACCESSION

VERSION AW091579

KEYWORDS

SOURCE EST.

ORGANISM Lycopersicon esculentum (Solanum lycopersicum)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

1 (bases 1 to 456)

AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J., Roming, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

TITLE

Location/Qualifiers

JOURNAL

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

FEATURES

5 prime sequence.

source

1..456

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="Rio Grande PtoR"

/db_xref="taxon:4081"

/clone="cLET15H6"

/tissue_type="leaf"

/dev_stage="4-6 week old plants"

/lab_host="X11-Blue MRF"

/clone_lib="tomato mixed elicitor, BTI"

/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; cLET: Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."

ORIGIN

source

1..612

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

Alignment Scores:

Pred. No.: 6,27e-24 Length: 456
 Score: 314.00 Matches: 83
 Percent Similarity: 59.7% Conservative: 12
 Best Local Similarity: 52.2% Mismatches: 40
 Query Match: 27.1% Indels: 24
 DB: 7 Gaps: 7

US-10-718-311-8 (1-227) x AW091579 (1-456)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 16 ATGGCTTCCTCTGTCTTCATTTCTTCAGCAGCTGTTCACACGACGACATGTTACACAGCT 75
 |||||
 QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 |||||
 Db 76 AGCATGGTGGACCTTTCACCTGTCTCAAAATCTTCAGCCACTTTCCTGTGTACAAAGAAG 135
 |||||
 QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 136 CAAACCTTGACATCACTTCCTTCATTTGCTAGCAATGGTGGAGAGTTAGTCGATGCAAGTG 195
 |||||
 QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
 |||||
 Db 196 TGGCCACCAATTAAC-----ATGAAGAGTACGAGACACTCTCATACCTTCCTGATTGTG 249
 |||||
 QY 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
 |||||
 Db 250 TCT-----GACGAGCAATTCCTAGTGAATTTAGTACCTTTTGAAAAATGATGGGTT 303
 |||||
 QY 99 LysArgPheGlnGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118
 |||||
 Db 304 CTTGCTGCAATTTGAG-----ACTGAGCAGGATTGTG 339
 |||||
 QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuProLysGluSerArgTyr 136
 |||||
 Db 340 TACCGTGAGAACCAACCAAGTCAACCAAGATACAT-----GATGGAAGGTAC 384
 |||||
 QY 137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
 |||||
 Db 385 TGGACCATGTGAAGTTGGCTATGTTGGTGTCATGTATGATCAACCAAGTGTGGCT 441
 |||||

AW622820 612 bp mRNA linear EST 18-MAY-2001
 EST306806 tomato flower buds 3-8 mm, Cornell University
 Lycopersicon esculentum cDNA clone cTOB1M21 5', mRNA sequence.

ACCESSION

VERSION AW622820

KEYWORDS

SOURCE EST.

ORGANISM Lycopersicon esculentum (Solanum lycopersicum)

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 612)

van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Roming, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.

Generation of ESTs from tomato flower tissue, 3-8 mm buds

Unpublished (1999)

Other ESTs: EST296223

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

Location/Qualifiers

1..612

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

```

/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT081W21"
/tissue_type="flower"
/dev_stages="3-8mm buds"
/clone_lib="tomato flower buds 3-8 mm, Cornell University"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9, 77e-24 Length: 612
Score: 314.00 Matches: 83
Percent Similarity: 59.7% Conservative: 12
Best Local Similarity: 52.2% Mismatches: 40
Query Match: 27.1% Indels: 24
DB: 7 Gaps: 7

```

US-10-718-311-8 (1-227) x AW622820 (1-612)

```

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 9 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGT 68

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 69 AGCATGGTGTGACCTTCTGCTGCTCAATCTTCAGCCACTTCCCTCTGTTACAAAGAAG 128

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 129 CAAACCTTGACATCCTCCATTCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 188

Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeu 80
Db 189 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCTCGAATTG 242

Qy 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db 243 TCT-----GACGACCAATGCTTAGTGAATGAGTACCTTTTGAAAAATGGATGGTT 296

Qy 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
Db 297 CCTTGCTTGGAAATTTGAG-----ACTGAGCAGCGATTGCTC 332

Qy 119 -----GluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 333 TACCGTGAGAACAAACAGTCACCGATCATAT-----GATGGAAGGTAC 377

Qy 137 TrpLeu-----ArgGluLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
Db 378 TGGACCATGTGGAAGTCGCTATGTTGGGTGCACTGATGCAACCAAGTGTGGCT 434

```

RESULT 9

AW041064

LOCUS

EST283928 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA

clone cLET11A11, mRNA sequence.

ACCESSION

AW041064

VERSION

AW041064.1

KEYWORDS

SOURCE

ORGANISM

Lycopersicon esculentum (Solanum lycopersicum)

Eukaryota; Viridiplantae;

Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 420)

REFERENCE

AUTHORS

D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,

TITLE

JOURNAL

COMMENT

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

Contact: CUGI

Clemson University

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

FEATURES

source

1..420

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="Rio Grande PtoR"

/db_xref="taxon:4081"

/clone="cLET11A11"

/tissue_type="leaf"

/dev_stages="4-6 week old plants"

/lab_hosts="XLI-Blue MRF"

/clone_lib="tomato mixed elicitor, BTI"

/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:

XhoI; cLET - Inoculated with a variety of disease response

elicitors. Plants exposed to 2,6 dichloroisocotinic

acid, BTH, jasmonic acid, ethylene, fenthion, EIX,

okadaic acid, or systemin prior to tissue harvest. EcoRI

site was destroyed during cloning."

ORIGIN

Alignment Scores:

```

Pred. No.: 6, 28e-24 Length: 420
Score: 313.50 Matches: 79
Percent Similarity: 63.8% Conservative: 11
Best Local Similarity: 56.0% Mismatches: 30
Query Match: 27.0% Indels: 21
DB: 7 Gaps: 6

```

US-10-718-311-8 (1-227) x AW041064 (1-420)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20

Db 19 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTTTCACACAGCAGCAATGTTACACAAGT 78

Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40

Db 79 AGCATGGTGTGACCTTCTGCTGCTCAATCTTCAGCCACTTCCCTCTGTTACAAAGAAG 138

Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

Db 139 CAAACCTTGACATCCTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 198

Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeu 80

Db 199 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCTCGAATTG 252

Qy 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98

Db 253 TCT-----GACGACCAATGCTTAGTGAATGAGTACCTTTTGAAAAATGGATGGTT 306

Qy 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118

Db 307 CCTTGCTTGGAAATTTGAG-----ACTGAGCAGCGATTGCTC 342

Qy 119 -----GluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136

Db 343 TACCGTGAGAACAAACAGTCACCAAGATCATAT-----GATGGAAGGTAC 387

Qy 137 Trp 137

Db 388 TGG 390

RESULT 10

AI775340

LOCUS

DEFINITION

EST256440 tomato resistant, Cornell Lycopersicon esculentum cDNA

```

clone cLER15A14, mRNA sequence.
AI775340
AI775340.1 GI:5273381
EST.
Lycopersicon esculentum (Solanum lycopersicum)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 422)
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. .422
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="cLER15A14"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/clone_lib="tomato resistant, Cornell"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
source
Alignment Scores:
Pred. No.: 6,33e-24 Length: 422
Score: 313.50 Matches: 79
Percent Similarity: 63.8% Conservative: 11
Best Local Similarity: 56.0% Mismatches: 30
Query Match: 27.0% Indels: 21
DB: 1 Gaps: 6

US-10-718-311-8 (1-227) x AI775340 (1-422)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 6 ATGGCTTCCTCTGTCATTTCTTTCCAGCAGCTGTGGCCACGCGCAGCAATGTTACACAGCT 65

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 66 AGCATGGTTGCACCTTTCTACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAG 125

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 126 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGGAAGAGTTAGCTGCATGCAGGTG 185

QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 186 TGCCACCACCAATTAAC-----ATGAAGAAGTAGCAGACACTCTCATPACCTTCCTGATTTG 239

QY 81 ProAlaLeuAspProGlnLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db 240 TCT-----GACGAGCAATTCCTTAGTGAAATGAGTACCTTCTGAAAAATGATGGGT 293

QY 99 LysArgPheGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
Db 294 CCTTGCTTGAATTTGAG-----ACTGAGCAGCAGGATTTGTC 329

```

```

QY 119 -----GluGlnAsnGluIleProGluGluProLeuLeuProLysGluSerArgTyr 136
Db 330 TACCGTGAGACACAAAGTCACGACTACTAT-----GATGAAGGTAC 374

QY 137 Trp 137
Db 375 TGG 377

RESULT 11
AW093612
LOCUS
DEFINITION
clone cLET25B3, mRNA sequence.
ACCESSION
AW093612
VERSION
AW093612.1 GI:6059207
KEYWORDS
EST.
SOURCE
Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 449)
D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato leaf tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
Location/Qualifiers
1. .449
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"
/clone="cLET25B3"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="XL1-Blue MRF,"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

ORIGIN
source
Alignment Scores:
Pred. No.: 7.89e-24 Length: 449
Score: 313.00 Matches: 84
Percent Similarity: 59.7% Conservative: 11
Best Local Similarity: 52.8% Mismatches: 40
Query Match: 27.0% Indels: 24
DB: 7 Gaps: 7

US-10-718-311-8 (1-227) x AW093612 (1-449)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 9 ATGGCTTCCTCTGTCATTTCTTTCCAGCAGCTGTGGCCACGCGCAGCAATGTTACACAGCT 68

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 69 AGCATGGTTGCACCTTTCTACTGGTCTCAAAATCTTCAGCCACTTTCCTCTGTTACAAAGAG 128

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60

```

```

Db      129 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCGGTG 188
Qy      61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db      189 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACACTCTCATACATCTCCCGATTG 242
Qy      81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db      243 TCT-----GAGCAGCAATTCCTAGTGAATGAGTACCTTTTGAAAAATGGATGGTT 296
Qy      99 LysArgPheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118
Db      297 CTTGCTTGGAAATTGAG-----ACTGACGACGGATTTGTC 332
Qy      119 -----GluGlnAnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db      333 TACCGTGAGAACAAAGTCACAGGATACAT-----GATGGAAGGTAC 377
Qy      137 TrpLeuArgGluIleLeuLeu-----CysAlaAspGlyGluProTrpLeuAla 152
Db      378 TGGACCATGNTGAAGTTCCTATGTTGGGTGCACTGATGCAACCAAGTGTGGCT 434

RESULT 12
AW622747 533 bp mRNA linear EST 18-MAY-2001
LOCUS EST306884 tomato flower buds 3-8 mm, Cornell University
DEFINITION Lycopersicon esculentum cDNA clone cTOB402 5', mRNA sequence.
ACCESSION AW622747
VERSION AW622747.1 GI:7335841
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 533)
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernan,W., Frazer,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato flower tissue, 3-8 mm buds
Unpublished (1999)
Contact: CUGI
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
1..533
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOB402"
/tissue_type="flower"
/dev_stage="3-8mm buds"
/clone_lib="tomato flower buds 3-8 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksey; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN
Alignment Scores:
Pred. No.: 1.02e-23 Length: 533
Score: 313.00 Matches: 83
Percent Similarity: 60.4% Conservative: 13
Best Local Similarity: 52.2% Mismatches: 39
Query Match: 27.0% Indels: 24
Db: 7 Gaps: 7

```

```

US-10-718-311-8 (1-227) x AW622747 (1-533)

Qy      1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAenValThrGlnAla 20
Db      17 ATGCTTCCTCTGTCACTTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 76
Qy      21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db      77 AGCATGGTTGCACCTTTCACCTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 136
Qy      41 GlnAnLeuAspIleThrSerIleAlaSerAnGlyGlyArgValSerCysMetGlnVal 60
Db      137 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCGGTG 196
Qy      61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db      197 TGGCCACCAATTAAC-----ATGAAGAAGTACGAGACACTCTCATACCTTCCTGATTG 250
Qy      81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
Db      251 TCT-----GACGAGCAATTCCTTAATCAAAATTGAGTACCTTTTGAAAAATGGATGGTT 304
Qy      99 LysArgPheGluGlnGlnGlyThrValSerValThrMetIleArgGluGlyPheVal 118
Db      305 CTTGCTTGGAAATTGAG-----ACTGACGACGGATTTGTC 340
Qy      119 -----GluGlnAnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db      341 TACCGTGAGAACAAAGTCACAGGATACAT-----GATGGAAGGTAC 385
Qy      137 TrpLeuArg-----GluIleLeuLeuCysAlaAspGlyGluProTrpLeuAla 152
Db      386 TGGACCATGTTGAAGTTCCTATGTTGGGTGCACTGATGCAACCAAGTGTGGCT 442

RESULT 13
BG129744 621 bp mRNA linear EST 31-JAN-2001
LOCUS EST475390 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION cTOF28E15 5' sequence, mRNA sequence.
ACCESSION BG129744
VERSION BG129744.1 GI:12629932
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 621)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
Contact: CUGI
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..621
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF28E15"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

FEATURES
source
1..621
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF28E15"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/clone_lib="tomato shoot/meristem"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."

ORIGIN

```

Alignment Scores: 1.29e-23 Length: 621
 Pred. No.: 313.00 Matches: 94
 Score: 313.00 Matches: 94
 Percent Similarity: 50.1% Conservative: 15
 Best Local Similarity: 43.1% Mismatches: 61
 Query Match: 27.0% Indels: 48
 DB: 2 Gaps: 9

US-10-718-311-8 (1-227) x BG129744 (1-621)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 15 ATGGCTCTCTGCTGCTATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAGCT 74
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 |||||
 Db 75 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 134
 |||||

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 135 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 194
 |||||

QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
 |||||
 Db 195 TGGCCACCAATTAAC-----ATGAAGAAGTAGCAGACACTCTCATACCTTCCTGATTGG 248
 |||||

QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
 |||||
 Db 249 TCT-----GACGAGCAATGCTTAGTGAAATGAGTACTTTTGAATAATGATGGTT 302
 |||||

QY 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
 |||||
 Db 303 CCTTGTCTGGATTTCAG-----ACTGAGCAGCGATTGTC 338
 |||||

QY 119 -----GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
 |||||
 Db 339 TACCGTGTGAGAACCAAGTCACCAAGGATACTAT-----GATGGAAGGTAC 383
 |||||

QY 137 TrpLeu-----138
 |||||

Db 384 TGGACCATGTGGGAAGTTGCCTATGTTGGTGCACTGATGCAACCAAGTGTGCTGA 443
 |||||

QY 139 -----ArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeu-----151
 |||||

Db 444 GGTTCAGAGCGCTAAAGGACATACCCCAAGCATGGGTGAGCAATCATTCGATTCGACAA 503
 |||||

QY 152 ---AlaGlyArgThrValProValSerThrLeuSerGlyProGluLeuAlaLeuGln 170
 |||||

Db 504 TGTGCGTCAAGTCAGCTGATCATGTTTCATGCTACAGCCAGAGCGCTACTAAGTTTC 563
 |||||

QY 171 LysLeuGly---LysThrProLeuGlyArgTyrLeuPheThrSerSerThrLeu 187
 |||||
 Db 564 ATATTAGGACAAATGCTTTAGGACACTTTGTTTAAATGCTACTTTA 617
 |||||

RESULT 14
 AW094716 416 bp mRNA linear EST 18-MAY-2001
 LOCUS EST287896 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION
 AW094716
 VERSION
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 416)
 D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J.,
 Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W.,
 Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and
 Giovannoni, J.

TITLE Generation of ESTs from tomato leaf tissue
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 5 prime sequence.

FEATURES
 Location/Qualifiers
 1..416
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande PtoR"
 /db_xref="taxon:4081"
 /clone="cLET29N8"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XL1-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, Elix,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

ORIGIN
 Alignment Scores: 7.98e-24 Length: 416
 Pred. No.: 312.50 Matches: 78
 Score: 312.50 Matches: 78
 Percent Similarity: 65.5% Conservative: 13
 Best Local Similarity: 56.1% Mismatches: 31
 Query Match: 27.0% Indels: 17
 DB: 7 Gaps: 5

US-10-718-311-8 (1-227) x AW094716 (1-416)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
 |||||
 Db 6 ATGGCTCTCTGCTGCTATTTCTTCAGCAGCTGTGGCCACGAGCAATGTTACACAGCT 65
 |||||

QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
 |||||
 Db 66 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 125
 |||||

QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
 |||||
 Db 126 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 185
 |||||

QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
 |||||
 Db 186 TGGCCACCAATTAAC-----ATGAAGAAGTAGCAGACACTCTCATACCTTCCTGATTGG 239
 |||||

QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeu-----LeuLeuGluAspSerMetThr 98
 |||||
 Db 240 TCT-----GACGAGCAATGCTTAGTGAAATGAGTACTTTTGAATAATGATGGTT 293
 |||||

QY 99 LysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheVal 118
 |||||
 Db 294 CCTTGTCTGGAAATTTGAGACTGACGACGAGTCTGTC-----TACCGT 335
 |||||

QY 119 GluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrp 137
 |||||
 Db 336 GAGAACCAAGTCACCCAGACTACTAT-----GATGAAGGTACTGG 377
 |||||

RESULT 15
 BP896321 434 bp mRNA linear EST 05-SEP-2005
 LOCUS BP896321 Lycopersicon esculentum leaf Lycopersicon esculentum cDNA
 DEFINITION
 clone LA08CH03 5', mRNA sequence.
 ACCESSION BP896321
 VERSION BP896321
 KEYWORDS BP896321.1 GI:58239792
 EST.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 11, 2006, 19:35:08 ; Search time 631.516 Seconds
(without alignments)
3759.291 Million cell updates/sec

Title: US-10-718-311-8
Perfect score: 1159
Sequence: 1 MASSVISNAVATRNVTQA.....RLSGKPLLLTELFLPASPLY 227

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5244920 seqs, 3486124231 residues
Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/abs/ABSSWEB spool/US10718311/runat_11072006_111518_5144/app_query.fasta.1
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNIT5=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss05p
-USER=US10718311@CGN_1_1_1423 @runat_11072006_111518_5144 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 8.*

```
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*
15: Geneseqn2006s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	1159	100.0	684	6	ABA91838 Chloropla
2	1159	100.0	684	6	ABK47712 DNA seque
3	1159	100.0	684	10	ADC25988 Tomato/Es

4	1159	100.0	684	14	ADV91622	Adv91622 Tomato ru
5	862	74.4	3452	14	ADV91636	Adv91636 C. paraps
6	856	73.9	854	4	AAD06956	AAD06956 ubiC gene
7	856	73.9	6641	4	AAD06957	AAD06957 pME2 comp
8	854	73.7	495	6	AAQ92409	AAQ92409 E.coli ch
9	854	73.7	495	6	ABA91837	ABA91837 Escherich
10	854	73.7	495	6	ABK47709	ABK47709 DNA seque
11	854	73.7	495	10	ADC25984	ADC25984 Escherich
12	854	73.7	498	14	ADV91618	Adv91618 Escherich
13	853	73.6	2000	2	AAT29820	AAT29820 E. coli u
14	851	73.4	1971	14	ADV91643	Adv91643 C. paraps
15	813	70.1	1207	5	AAS90335	AAS90335 DNA encod
16	795	68.6	1218	2	AAQ67671	AAQ67671 E.coli ub
17	673	58.1	645	11	ACH95439	ACH95439 Klebsiell
18	417	36.0	519	10	ACF70957	ACF70957 Phototrab
19	417	36.0	110000	10	ACF67367	ACF67367 40
20	417	36.0	110000	10	ACF65388	ACF65388 07
21	401	34.6	537	10	ADF02837	ADF02837 Bacterial
22	282.5	24.4	615	10	ADC76949	ADC76949 DNA homol
23	281.5	24.3	608	10	ADC76944	ADC76944 DNA homol
24	281.5	24.3	684	10	ADC76948	ADC76948 DNA homol
25	281.5	24.3	859	11	ADM45065	ADM45065 Insect re
26	277.5	23.9	492	10	ADC76164	ADC76164 DNA homol
27	277.5	23.9	499	6	AD35635	AD35635 Nicotiana
28	277.5	23.9	504	10	ADK59824	ADK59824 Plant DNA
29	277.5	23.9	619	11	ADM45158	ADM45158 Insect re
30	276.5	23.9	489	10	ADK56120	ADK56120 Plant DNA
31	276.5	23.9	581	10	ADC75089	ADC75089 N bentham
32	276.5	23.9	609	10	ADC76965	ADC76965 DNA homol
33	276.5	23.9	614	10	ADC76956	ADC76956 DNA homol
34	276.5	23.9	632	10	ADC76953	ADC76953 DNA homol
35	276.5	23.9	718	10	ADK58382	ADK58382 Plant DNA
36	276.5	23.9	736	10	ADK54321	ADK54321 Plant DNA
37	276.5	23.9	736	10	ADK57660	ADK57660 Plant DNA
38	276.5	23.9	841	11	ADM44879	ADM44879 Insect re
39	276.5	23.9	847	11	ADM45447	ADM45447 Insect re
40	269	23.2	204	6	AAI70688	AAI70688 Tobacco R
41	269	23.2	297	10	ADK59826	ADK59826 Plant DNA
42	269	23.2	377	10	ADK56121	ADK56121 Plant DNA
43	269	23.2	924	10	ADC75566	ADC75566 DNA homol
44	269	23.2	958	10	ADC76165	ADC76165 DNA homol
45	266	23.0	668	10	ADK59825	ADK59825 Plant DNA

ALIGNMENTS

RESULT 1
ABA91838
ID ABA91838 standard; DNA; 684 BP.
XX ABA91838;
AC ABA91838;
XX 29-AUG-2003 (revised)
DT 15-MAY-2002 (first entry)
XX Chloroplast transit peptide-chorismate pyruvate lyase fusion gene.
XX Chloroplast transit peptide; tomato; Rubisco; plant;
KW ribulose-1,5-bisphosphate carboxylase; chorismate pyruvate lyase; CPL;
KW enzyme; p-hydroxybenzoic acid; transgenic plant; gene; ds.
XX Lycopersicon esculentum.
OS Escherichia coli.
OS Chimeric.
XX Key Location/Qualifiers
FT CDS 1..495
FT /*tag= a
FT /*product= "chloroplast-targeted CPL fusion"
FT transit_peptide 1..186
FT /*tag= b
FT /*note= "tomato Rubisco chloroplast targeting sequence"
FT mat_peptide 187..492

```

FT      /*tag= c
FT      /note= "E. coli chorismate pyruvate lyase"
PN      WO200194607-A2.
XX      13-DEC-2001.
XX
XX      22-MAY-2001; 2001WO-US016661.
XX
XX      02-JUN-2000; 2000US-0209854P.
XX
XX      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX      Meyer K, Van Dyk DE, Viitanen PV;
XX
XX      WPI; 2002-226795/28.
XX      P-PSDB; AAM50959.
XX
XX      Producing para-hydroxy benzoic acid in green plant, comprises expression
XX      of unique expression cassette containing gene encoding chorismate
XX      pyruvate lyase operably linked to specific chloroplast targeting
XX      sequence.
XX
XX      Claim 17; Page 57; 60pp; English.
XX
XX      The present nucleotide sequence is that of an open reading frame encoding
XX      chloroplast-targeted chorismate pyruvate lyase (CPL) fusion protein TP-
XX      CPL (see AAM50959). PCR (see ABA91841-42) was used to generate a DNA
XX      fragment corresponding to the transit peptide of tomato Rubisco small
XX      subunit and the first 4 amino acid residues of mature Rubisco. The DNA
XX      fragment was ligated into pET24a-CPL, which carries the Escherichia coli
XX      CPL open reading frame (see ABA91837). The TP-CPL construct is an example
XX      of expression cassettes of the invention that are designed for the high-
XX      level production of p-hydroxybenzoic acid (pHBA) in higher plants.
XX      Cleavage of TP-CPL in the chloroplast releases a novel polypeptide (see
XX      AAM50961) that has full enzyme activity, converting chorismate to pHBA. A
XX      claimed plant comprising a CPL expression cassette is selected from
XX      soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
XX      barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,
XX      millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric
XX      component of liquid crystal polymers which have application in the
XX      automotive, electrical and other industries. (Updated on 29-AUG-2003 to
XX      standardize OS field)
XX
XX      SQ      Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      5.77e-119      Length:      684
Score:      1159.00      Matches:      227
Percent Similarity:      100.0%      Conservative:      0
Best Local Similarity:      100.0%      Mismatches:      0
Query Match:      100.0%      Indels:      0
DB:      6      Gaps:      0

US-10-718-311-8 (1-227) x ABA91838 (1-684)

Qy      1      MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db      1      ATGGCTTCCTCTGTCTATTTCTTCAGCAGCTGTTCACACGACGACGACGATGTTACACAGCT 60

Qy      21      SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db      61      AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCTGTTTACAAAGAAG 120

Qy      41      GlnAsnLeuAspIleThrSerIleAlaSerAndGlyGlyArgValSerCysMetGlnVal 60
Db      121      CAAACCTTGACATCACATCTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180

Qy      61      TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db      181      TGGCATATGTACACACCCCGCGTTAAGCAACTGCGTGGCTGGCTATTGTAAGAGATC 240

Qy      81      ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100

```

```

Db      241      CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAGCT 300
Qy      101      PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db      301      TTTGAACAGCAGCAGGAAACCGTTAAGCGTGACGATGATCCCGAAGGGTTTGTGAGCAG 360
Qy      121      AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db      361      AATGAAATCCCGAAGAACTGCCCGTCTGCTCCGAAAGAGTCTCGTTACTGTGTACGTGAA 420
Qy      141      IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160
Db      421      ATTTTGTATGTGCGGATGGTGAACCGTGGCTTCCCGTTCGTACCTGTTCTCTGTGTCA 480
Qy      161      ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db      481      ACGTTAAGCGGCGCGAGCTGGCGTTACAAAATTGGGTAAACGCCGTTTAGGACGCTAT 540
Qy      181      LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db      541      CTGTTTCATCATCATCGACATTAAACCCGGACTTTTATTGAGATAGGCCCGGTGATGCCG 600
Qy      201      TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db      601      TGGGGCGGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy      221      LeuProAlaSerProLeuTyr 227
Db      661      TTACCGGCGTCACCGTTGTATC 681

RESULT 2
ABK47712
ID      ABK47712 standard; DNA; 684 BP.
XX
XX      ABK47712;
XX
XX      18-JUN-2002 (first entry)
XX
XX      DNA sequence of ORF for TP-CPL fusion protein.
XX
XX      SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester; plant;
XX      sinapoylglucose:maltate sinapoyltransferase; glycosylated aromatic acid;
XX      malate conjugated aromatic acid; polymer synthesis; ubiC; TP-CPL; gene;
XX      carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;
XX      tomato Rubisco small subunit precursor; rbcS2; mutant; ds.
XX
XX      Escherichia coli.
XX      Lycopersicon esculentum.
XX      Synthetic.
XX      Chimeric.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..684
XX              /*tag= a
XX              /product= "TP-CPL fusion protein"

WO200204653-A2.
17-JAN-2002.
05-JUL-2001; 2001WO-US021283.
07-JUL-2000; 2000US-0216615P.
(DUPO ) DU PONT DE NEMOURS & CO E I.
Flint D, Meyer K, Viitanen PV;
WPI; 2002-303779/34.
P-PSDB; AAU77942.
Producing aromatic acid conjugates, involves contacting glycosylated

```

PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the
PT presence of sinapoylglucose:malate sinapoyltransferase.

PS Example; Page 70; 72pp; English.

XX The present invention relates to the isolation of Arabidopsis thaliana
CC gene (SNG1, sinapoylglucose accumulator 1) encoding
CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the
CC substitution of a glucose moiety on a glycosylated aromatic acid with a
CC malate moiety to form a malate conjugated aromatic acid. The enzyme is
CC useful for producing malate conjugated aromatic acids, carboxylic acid
CC conjugated aromatic acids or aromatic esters. The malate conjugated
CC aromatic acids are useful in the synthesis of various polymers. The
CC present sequence representing the ORF (open reading frame) for TP-CPL
CC (tomato Rubisco small subunit precursor for rbcS2-Escherichia coli
CC chorismate pyruvate lyase (CPL)) fusion protein is described in the
XX examples of the present invention

SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,77e-119 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-718-311-8 (1-227) x ABK47712 (1-684)

QY 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
DB 1 ATGGCTTCCTCTGTCAATTTCTTCAGCAGCTGTTCACACGCGCAGCAATGTTACACAAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
DB 61 AGCATGGTTGCACCTTTCCTGGTCTCAATCTTCAGCCACCTTTCCTCTGTACAAAGAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
DB 121 CAAACACCTTGACATCACTTCCATTTGCTAGCAATGGTGGAAAGATTAGCTGCATGCAGGTG 180
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
DB 181 TGGCATATGTACACCCCGGTTAACGCAACTGCGTGGCTGCTGCTATTTGTAAGAGATC 240
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
DB 241 CTGCGCTGGATCCGCAACTGCTGACTGGCTGGTGGAGGATTCATGACNAACGT 300
QY 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
DB 301 TTTGAACAGCAGGAAACCGGTAAAGCGTGAAGCGTGAAGCGTGGTGGTGGTGGTGGTGG 360
QY 121 AsnGluIleProGluLeuLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
DB 361 AATGAATCCCCGAAAGACTGCGCTGCGCGAAAGAGTCTGTTACTGGTTACGTGAA 420
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160
DB 421 ATTTTGTATGTCCGATGTTGACCGTGGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 480
QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
DB 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTGGGTAAACCGCGTTAGACGCTAT 540
QY 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
DB 541 CTGTTACATCATCGACATTAACCCGGGACTTTTATTGAGATAGGCGCGTATGCCGGCTG 600
QY 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
DB 601 TGGGGCGGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTTGCTTAACAGACTGTTT 660

QY 221 LeuProAlaSerProLeuTyr 227
DB 661 TTACCGGCGTCCACCGTTGTAC 681

RESULT 3

ADC25988

ID ADC25988 standard; DNA; 684 BP.

XX

AC ADC25988;

XX 18-DEC-2003 (first entry)

XX Tomato/Escherichia coli chimeric TP-CPL DNA.

XX UDP-glucosyltransferase; pHBA; p-hydroxybenzoic acid ester glucoside;

KW liquid crystal polymer; LCP; methylparaben; preservative; food;

KW cosmetic industry; ds; gene; tomato; TP-CPL; chimeric;

KW chorismate pyruvate lyase; plant.

XX Chimeric.

OS Lycopersicon esculentum.

OS Escherichia coli.

XX Key Location/Qualifiers

FT 1..684

FT /product= a

FT /tag= a

FT protein

XX WO2003066836-A2.

XX 14-AUG-2003.

XX 06-FEB-2003; 2003WO-US005863.

XX 07-FEB-2002; 2002US-0355511P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;

XX WPI: 2003-767259/72.

XX P-PSDB; ADC25989.

PT New nucleic acid encoding UDP-glucosyltransferase, useful for preparing
PT cells that produce p-hydroxybenzoic acid glucose ester, also the new
PT enzymes.

PS Example 9; SEQ ID NO 41; 161pp; English.

XX The invention relates to a novel isolated nucleic acid that encodes a UDP
-glucosyltransferase. The method of the invention may be used to
transform microorganisms or green plant cells so that these produce a
higher level of high-volume chemicals or materials, such as pHBA (p-
hydroxybenzoic acid) ester glucoside which is an intermediate for pHBA, a
monomer for liquid crystal polymers (LCPs) and starting material for
methylparaben. Methylparaben is a preservative commonly used in the food
and cosmetic industries. The encoded enzymes may be used for in vitro
production of these compounds and for identifying similar enzymes by
sequence comparison. The current sequence is that of the
tomato/Escherichia coli chimeric TP-CPL (chorismate pyruvate lyase) DNA
of the invention.

SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,77e-119 Length: 684
Score: 1159.00 Matches: 227
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

```
US-10-718-311-8 (1-227) x ADC25988 (1-684)
QY 1 MetAlaSerSerValIleSerSerAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCTATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
QY 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTGCACCTTCTACTGGTCTCAATCTTCAGCCACTTTCCTCTGTTACAAAGAG 120
QY 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAAACCTTGACATCACTCCATTCCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
QY 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTTCACCCCGCGTTAACGCCAACTGCGTGCCTGCTATTTGTAAGAGATC 240
QY 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGTGCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAGCT 300
QY 101 PheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGln 120
Db 301 TTTGAACAGCAGGAGAAACGGTAAGCGGTGACGATGATCCGCCAAGGGTTTGTGAGCAG 360
QY 121 AsnGluIleProGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAAATCCCGAAGAACTGCCGCTGCTGCCGAAGAGTCTCTGTTACTGGTTACGTGAA 420
QY 141 IleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSer 160
Db 421 ATTTTGTATGTGCCGATGGTGAACCGTGGCTTGGCGGTGTACCGTGTCTCTGTGTCA 480
QY 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCCGCTAGGACGCTAT 540
QY 181 LeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTTCATCATCATCGACATTAACCGGACATTTATTGAGATAGGCGGTGATGCCGGGCTG 600
QY 201 TrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGGCGAGCGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGACTGTTT 660
QY 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGGCTCACCGCTTGTAC 681
RESULT 4
ADV91622
ID ADV91622 standard; DNA; 684 BP.
XX
AC ADV91622;
XX
DT 10-MAR-2005 (first entry)
XX
DE Tomato rubisco small subunit precursor TP-E. coli CPL chimeric DNA.
XX
KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;
KW genetically engineered microorganism; antioxidant; antimicrobial agent;
KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
KW neoplasm; PHBA; para-hydroxybenzoic acid; p-hydroxybenzoic acid;
KW rubisco small subunit precursor; transit peptide;
KW chloroplast transit peptide; TP; chloroplast; gene; ds.
XX
OS Lycopersicon esculentum.
OS Escherichia coli.
OS Chimeric.
```


Query Match:	73.9%	Indels:
DB:	4	Gaps:
US-110-718-311-8 (1-227) x AAD06956 (1-854)		
Qy	57	CysMetClnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
Db	303	TGTAACGGAGAGTTCGGCATGTACACCCCGCGTTAACGCAACTCGCTGCGCTGCGCTAT 362
Qy	77	CysLysGluLeuProAlaLeuAAspProGlnLeuLeuAspTfPLeuLeuGluAAspSer 96
Db	363	TGTAAGAGAGATCCCTGCGCTCGATCCGCAACTGTCTCGACTGGCTGTGCTGGAGGATTCC 422
Qy	97	MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db	423	ATGACAAAAACGTTTTGAACACAGCAGGGAATAACCGGTAAACGCTGACGATATCCGCGAAGG 482
Qy	117	PheValGluGlnAenGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
Db	483	TTTGTGCGAGCAGATGAATCCCGAAGAACTGCCGTGCTGCCGAAGAAGTCTCGTTAC 542
Qy	137	TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db	543	TGTTACGTAGTAAATTTTTGTATGTCGCATGTGTGAACCGTGGCTTCCCGCTCGTACCGTC 602
Qy	157	ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db	603	GTTCTCTGTCAACGTTAAGGGGCCGGAGCTGGCGTTACAAAAATTGGGTAAAAACGCG 662
Qy	177	LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
Db	663	TTAGGACGCTATCTCTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCCGT 722
Qy	197	AspAlaGlyLeuTrpGlyArgArgSerArgLeuArgGluSerGlyLysProLeuLeu 216
Db	723	GATGCCGGGCTGTGGGGGCGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTA 782
Qy	217	ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db	783	ACAGAACTGTTTTTACCGCGCTCACCGTTGTATC 815
RESULT 7		
AAD06957		
XX	AAD06957 standard; DNA; 6641 BP.	
XX	AAD06957;	
XX	06-AUG-2001 (first entry)	
XX	pME2 comprising ubiC, aroG, tktA and ppsA genes of the aromatic pathway.	
KW	p-hydroxybenzoic acid; pH5; aromatic pathway; carbon source; ubiC;	
KW	chorismate pyruvate lyase; ppsA; phosphoenolpyruvate synthase;	
KW	ppp synthase; tktA; transketolase; aroC; chorismate synthase; aroA;	
KW	5-enolpyruvylshikimate-3-phosphate synthase; EPSP synthase; aroL;	
KW	shikimate kinase II; aroB; 3-dehydroquinate synthase; DHQ synthase; aroG;	
KW	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DHAP synthase;	
XX	pME2; ds.	
OS	Escherichia coli.	
XX		
PN	US6210937-B1.	
XX		
PD	03-APR-2001.	
XX		
PF	22-APR-1998; 98US-00064693.	
XX		
PR	22-APR-1997; 97US-0044094P.	
XX		
PA	(BECH-) BECHTEL BWXT IDAHO LLC.	
XX		
PI	Ward TE, Watkins CS, Bulmer DK, Johnson BF, Amaratunga M;	
XX		

DR WPI; 2001-280857/29.
 XX Producing aromatic compounds, especially commercially acceptable levels
 PT of p-hydroxybenzoic acid, comprises developing genetically engineered
 PT bacteria that carry selected genes of the common aromatic pathway.
 XX
 XX
 PS Claim 4; Col 31-40; 25pp; English.
 XX
 XX The present invention relates to a method of producing p-hydroxybenzoic
 CC acid (pHB) which comprises transforming microorganisms with plasmids
 CC carrying selected genes of the common aromatic pathway and by conversion
 CC of carbon sources. The selected genes of the common aromatic pathway of
 CC the invention includes ubiC gene encoding chorismate pyruvate lyase, aroG
 CC structural gene encoding 3-deoxy-D-arabino-heptulosonate-7-phosphate
 CC (DHAP) isoenzyme synthase (phe), tktA gene encoding transketolase, ppsA
 CC gene encoding phosphoenolpyruvate (PEP) synthase, aroC gene encoding
 CC chorismate synthase, aroL gene encoding shikimate kinase II, aroA gene
 CC encoding 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase and aroB
 CC gene encoding 3-dehydroquinate (DHQ) synthase. The method is used for
 CC converting chorismate to a selected aromatic compound. The method is
 CC particularly used for the production of commercially acceptable levels of
 CC aromatic compounds, particularly p-hydroxybenzoic acid. The present
 CC sequence is pHE2 plasmid which comprises ubiC, aroG, tktA and ppsA genes
 CC involved in the aromatic pathway
 XX
 SQ Sequence 6641 BP; 1567 A; 1698 C; 1853 G; 1523 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,89e-84 Length: 6641
 Score: 856.00 Matches: 166
 Percent Similarity: 97.7% Conservatives: 1
 Best Local Similarity: 97.1% Mismatches: 4
 Query Match: 73.9% Indels: 0
 DB: 4 Gaps: 0

US-10-718-311-8 (1-227) x AAD06957 (1-6641)

QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
 DB 303 TGTAAAGAGATCCCTGCCCTCGATGTCACACCCCGGTAAACGCACTGGCTGCGCTAT 362
 QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
 DB 363 TGTAAAGAGATCCCTGCCCTCGATGTCACACCCCGGTAAACGCACTGGCTGCGCTAT 422
 QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
 DB 423 ATGACAAAACGTTTGAACAGCAGCGGAAAAACGGTAAGCGTGACGATGATCCGGAAGG 482
 QY 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
 DB 483 TTTGTGACGAGAATGAAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTAC 542
 QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
 DB 543 TGGTTACGTGAAATTTTGTATGTGCGGATGTGAACCGTGCTGCCGCTGCTACCGTC 602
 QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
 DB 603 GTTCTGTGTCAACGTTAAGCGCGCGAGCTGGCGTTACAAAAATTTGGGTAAACCGCG 662
 QY 177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
 DB 663 TTAGACGCTATCTGTTCATCATCATCAATTAACCCGGGACTTTATTGAGATAGCCGT 722
 QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
 DB 723 GATGCCGGCTGTGGGGGCGACGTTCCCGCTCGGATTAAACCGGTAAACCGCTGTGCTA 782
 QY 217 ThrGluLeuPheLeuProLysProLeuSerProLeuTyr 227
 DB 783 ACAGAACTGTTTACCGCGCTCACCGTTGTAC 815

RESULT 8

AAQ92409
 ID AAQ92409 standard; DNA; 495 BP.
 XX
 XX AAQ92409;
 XX
 DT 11-JAN-1996 (first entry)
 XX
 DE E.coli chorismate-pyruvate lyase coding sequence.
 XX
 KW Chorismate-pyruvate lyase gene; ubiC gene; transgenic plant;
 KW tobacco mosaic virus resistance; antiviral activity;
 KW p-hydroxybenzoic acid; ss.
 XX
 OS Escherichia coli.
 XX
 PN DE4423022-Cl.
 XX
 PD 24-MAY-1995.
 XX
 PF 30-JUN-1994; 94DE-04423022.
 XX
 PR 30-JUN-1994; 94DE-04423022.
 XX
 PA (HEID/) HEIDE L.
 XX
 PI Heide L, Siebert M, Severin K;
 XX
 DR WPI; 1995-186908/25.
 DR P-PSDB; AAR74742.
 XX
 PT Transgenic plants with increased content of resistance factor - contg.
 PT bacterial DNA coding for enzyme catalysing prodn. of resistance factor,
 PT e.g. antiviral p-hydroxy-benzoic acid.
 XX
 PS Claim 3; Page 5; 7pp; German.
 XX
 CC Transgenic plants containing a bacterial gene which encodes an enzyme
 CC able to catalyse prodn. of an antiviral, bactericidal, fungicidal or
 CC insecticidal factor are new. A preferred gene is ubiC from E.coli
 CC (AAQ92409) which codes for chorismate-pyruvate lyase (AAR74742). The
 CC lyase catalyses conversion of chorismate to p-hydroxybenzoic acid and
 CC transgenic tobacco plants which express the active enzyme are resistant
 CC to tobacco mosaic virus
 XX
 SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,04e-85 Length: 495
 Score: 854.00 Matches: 165
 Percent Similarity: 100.0% Conservatives: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 73.7% Indels: 0
 DB: 2 Gaps: 0
 US-10-718-311-8 (1-227) x AAQ92409 (1-495)
 QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82
 DB 1 ATGTCACACCCCGGTAAACGCACTGGCTGCGCTATTTGTAAGAGATCCCTGCC 60
 QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArgPheGlu 102
 DB 61 CTGGATCCGCAACTGCTGACCTGGCTGTGTCTGGAGATTCCATGACAAACCTTTGAA 120
 QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
 DB 121 CAGCAGGGGAAAAACGGTAAGCGTGACGATGATCCCGAAGGGTTTCTCCGACAGATGAA 180
 QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluLeuLeu 142
 DB 181 ATCCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGTTAGCTGAAATTTTG 240

QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
 DB 241 TTATGTGGCGATGGTAAACCGTGGTGGCGGTGCTACCGTCTCTGTGTCAACGTTA 300
 QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
 DB 301 AGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAACCGCGCTTAGGACGCTATCTGTTC 360
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArgAspAlaGlyLeuTrpGly 202
 DB 361 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGCGCGGTGATGCCGGCTGTGGGG 420
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
 DB 421 CGACGTTCCCGCTGGATTAGCGGTAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
 QY 223 AlaSerProLeuTyr 227
 DB 481 GCGTCACCGTTGTAC 495
 RESULT 9
 ABA91837
 ID ABA91837 standard; DNA; 495 BP.
 XX
 AC ABA91837;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Escherichia coli chorismate pyruvate lyase gene coding region.
 XX
 KW Chorismate pyruvate lyase; CPL; enzyme; p-hydroxybenzoic acid;
 KW transgenic plant; ubiC; gene; ds.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..495
 FT /*tag= a
 FT /partial
 FT /product= "Chorismate pyruvate lyase"
 FT /note= "the CDS does not include a stop codon"
 XX
 PN WO200194607-A2.
 XX
 XX 13-DEC-2001.
 XX
 PF 22-MAY-2001; 2001WO-US016661.
 XX
 XX 02-JUN-2000; 2000US-0209854P.
 XX
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Meyer K, Van Dyk DE, Viitanen PV;
 XX
 DR WPI; 2002-226795/28.
 XX
 DR P-PSDB; AAM50958.
 XX
 PT Producing para-hydroxy benzoic acid in green plant, comprises expression
 PT of unique expression cassette containing gene encoding chorismate
 PT pyruvate lyase operably linked to specific chloroplast targeting
 PT sequence.
 XX
 XX Claim 4; Page 55; 60pp; English.
 PS
 CC The present sequence is that of the coding region of the Escherichia coli
 CC strain W3110 ubiC gene in expression construct pET24a. The gene encodes
 CC chorismate pyruvate lyase (CPL, see AAM50958). It was obtained by PCR
 CC amplification of strain W3110 genomic DNA using primers (see ABA91839-40)
 CC based on the published E. coli ubiC gene. The CPL open reading frame is
 CC used in an expression cassette designed for the high-level production of
 CC p-hydroxybenzoic acid (pHBA) in green plants. The expression cassette
 CC comprises the CPL coding sequence operably linked to a promoter capable
 CC of driving protein expression in higher plants. The cassette also has a

CC sequence encoding a chloroplast transit peptide, its natural cleavage
 CC site, and a small portion of a transit peptide donor protein fused to the
 CC N-terminus of CPL. The chloroplast targeting sequence targets the foreign
 CC protein to the chloroplast and aids in its uptake into the organelle. The
 CC cleavage site is unique to the transit peptide, and cleavage of the
 CC chimeric protein encoded by the cassette at this site releases a novel
 CC polypeptide that has full enzyme activity, comprising the mature CPL
 CC enzyme and a small portion of transit peptide donor (see AAM50961). A
 CC plant comprising the CPL expression cassette is claimed, and may be
 CC soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
 CC barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,
 CC millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric
 CC component of liquid crystal polymers which have application in the
 CC automotive, electrical and other industries
 XX
 SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. NO.: 3.04e-85 Length: 495
 Score: 854.00 Matches: 165
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 73.7% Indels: 0
 DB: 6 Gaps: 0
 US-10-718-311-8 (1-227) x ABA91837 (1-495)
 QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluLeuProAla 82
 DB 1 ATGTACACCCCGCTTAACGCAACTCGTGGCTGCTATTGTAAAGAGATCCCTGCC 60
 QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrLysArgPheGlu 102
 DB 61 CTGGATCCGCACTGCTCGACTGGCTGTGGAGGATCCATGACAAACGTTTGAA 120
 QY 103 GlnGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGlnAsnGlu 122
 DB 121 CAGCAGGGAACCGTAAAGCGTGCATGATCGCGAAGGGTTTGTGAGCAGAGTAA 180
 QY 123 IleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGluLeu 142
 DB 181 ATCCCCGAAGAACTGCCGCTGCTCCGAAGAGTCTCGTTACTGTTACGTGAAATTTG 240
 QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
 DB 241 TTATGTCCGATGGTGAACCGTGGCTTGGCGGTGCTACCGTCTCTGTCTCAACGTTA 300
 QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
 DB 301 AGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAACCGCGCTTAGGACGCTATCTGTTC 360
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArgAspAlaGlyLeuTrpGly 202
 DB 361 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGCGCGGTGATGCCGGCTGTGGGG 420
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
 DB 421 CGACGTTCCCGCTGGATTAAAGCGGTAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
 QY 223 AlaSerProLeuTyr 227
 DB 481 GCGTCACCGTTGTAC 495
 RESULT 10
 ABAK47709
 ID ABAK47709 standard; DNA; 495 BP.
 XX
 AC ABAK47709;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE DNA sequence of ORF for CPL in pET24a E. coli expression construct.
 XX

KW SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester;
 KW sinapoylglucose:malate sinapoyltransferase; glycosylated aromatic acid;
 KW malate conjugated aromatic acid; polymer synthesis; ubiC; CPL; gene;
 KW carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;
 KW pET24a; ds.
 XX Escherichia coli.
 FH Key Location/Qualifiers
 FT CDS 1..495
 FT /*tag= a
 FT /partial
 FT /product= "CPL"
 FT /note= "This sequence lacks a stop codon"
 XX WO200204653-A2.
 PN 17-JAN-2002.
 XX 05-JUL-2001; 2001WO-US021283.
 XX 07-JUL-2000; 2000US-0216615P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Flint D, Meyer K, Viitanen PV;
 XX WPI; 2002-303779/34.
 XX P-PSDB; AAU77941.
 DR Producing aromatic acid conjugates, involves contacting glycosylated
 PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the
 PT presence of sinapoylglucose:malate sinapoyltransferase.
 XX Example; Page 68; 72pp; English.
 XX The present invention relates to the isolation of Arabidopsis thaliana
 CC gene (SNG1, sinapoylglucose accumulator 1) encoding
 CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the
 CC substitution of a glucose moiety on a glycosylated aromatic acid with a
 CC malate moiety to form a malate conjugated aromatic acid. The enzyme is
 CC useful for producing malate conjugated aromatic acids, carboxylic acid
 CC conjugated aromatic acids or aromatic esters. The malate conjugated
 CC aromatic acids are useful in the synthesis of various polymers. The
 CC present sequence representing the ORF (open reading frame) for chorismate
 CC pyruvate lyase (CPL) in pET24a Escherichia coli expression construct is
 CC described in the examples of the present invention
 XX SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,04e-85 Length: 495
 Score: 854.00 Matches: 165
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 73.7% Indels: 0
 DB: 6 Gaps: 0

US-10-718-311-8 (1-227) x ABK47709 (1-495)

QY 63 MetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIleProAla 82
 DB 1 ATGTACACCCCGGTTACGCAACTGCGTGGCGCTATTGTAAAGAGATCCCTGCC 60
 QY 83 LeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSerMetThrLysArgPheGlu 102
 DB 61 CTGGATCCCAACTGCTCACTGGCTGTCTGGAGGATCCATGACAAAGCTTTGAA 120
 QY 103 GlnGlnGlyThrValSerValThrMetIleArgGluGlyPheValGluGlnAenGlu 122
 DB 121 CAGCAGGGAAAAACGGTAGCGTACGATGATCCGCGAAGGGTTGTTCGAGCAGAGTAA 180
 QY 123 IleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIleLeu 142

DB 181 ATCCCGGAGAACTCCCGCTGCTCCGAAAGAGTCTCGTTACTGGTTACGTGAAATTTTG 240
 QY 143 LeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
 DB 241 TTATGTCCGATGGTGAACCGTGGCTTGCCGTCGTACCGTCTCTCTGTCACAGTTA 300
 QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeuPhe 182
 DB 301 AGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAAACGCCCTTAGGACGCTATCTGTTTC 360
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrpGly 202
 DB 361 ACATCATCGACATTAAACCGGACTTTATTGAGATAGGCCGCGTATGCCGGCTGTGGGG 420
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuLeuLeuLeuLeuPro 222
 DB 421 CGACGTTCCCGCTCGGATTAAACCGGTAACCGCTGTTGCTAACAGAACTGTTTTACCG 480
 QY 223 AlaSerProLeuTyr 227
 DB 481 GCGTCACCGTTGTAC 495

RESULT 11
 ADC25984
 ID ADC25984 standard; DNA; 495 BP.
 XX
 AC ADC25984;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Escherichia coli chorismate pyruvate lyase DNA.
 KW UDP-glucosyltransferase; PHBA; p-hydroxybenzoic acid ester glucoside;
 KW liquid crystal polymer; LCP; methylparaben; preservative; food;
 KW cosmetic industry; ds; gene; ubiC; CPL; chorismate pyruvate lyase.
 XX Escherichia coli.
 OS
 FH Key Location/Qualifiers
 CDS 1..495
 FT /*tag= a
 FT /partial
 FT /product= "Escherichia coli chorismate pyruvate lyase
 FT protein"
 FT /note= "No stop codon"
 XX WO2003066836-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 06-FEB-2003; 2003WO-US005863.
 XX
 PR 07-FEB-2002; 2002US-0355511P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Meyer K, Van Dyk DE, Viitanen PV;
 XX
 DR WPI; 2003-767259/72.
 XX
 DR P-PSDB; ADC25985.
 XX
 PT New nucleic acid encoding UDP-glucosyltransferase, useful for preparing
 PT cells that produce p-hydroxybenzoic acid glucose ester, also the new
 PT enzymes.
 XX
 PS Example 9; SEQ ID NO 37; 161pp; English.
 XX

CC The invention relates to a novel isolated nucleic acid that encodes a UDP
 CC -glucosyltransferase. The method of the invention may be used to
 CC transform microorganisms or green plant cells so that these produce a
 CC higher level of high-volume chemicals or materials, such as PHBA (p-
 CC hydroxybenzoic acid) ester glucoside which is an intermediate for PHBA, a

QY 143 LeuCyAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThrLeu 162
 Db 241 TTATGTGCCGATGCTGAACCGTGGCTTGGCGGTCTGACCGTCTGCTGTCTCAACGTTA 300
 QY 163 SerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgThrLeuPhe 192
 Db 301 AGCGGGCCGAGCTGGCGTTTACAAAATGGGTAAACGCCGTTAGACGCTATCTGTTT 360
 QY 183 ThrSerSerThrLeuThrArgAspPheIleGlyLeuGlyArgAspAlaGlyLeuTrpGly 202
 Db 361 ACATCATCGACATTAACCGGGACTTATTGAGATAGCGCTGATCGCGGCTGTGGGG 420
 QY 203 ArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeuPro 222
 Db 421 CGACGTTCCCGCTCGGATTAGCGGTAAACCGCTGTTGTTACAGAACTGTTTTTACCG 480
 QY 223 AlaSerProLeuTyr 227
 Db 481 CGGTACCGTTGTAC 495
 RESULT 13
 AAT29820
 ID AAT29820 standard; DNA; 2000 BP.
 AC AAT29820;
 XX
 XX
 DT 17-JAN-1997 (first entry)
 XX
 DE E. coli ubiquinone biosynthetic enzyme genes ubiC and ubiA.
 XX
 XX Ubiquinone-10; biosynthetic enzyme; ubiC; ubiA; gene;
 KW photosynthetic bacteria; recombinant production; plasmid pRSFAC;
 KW glutamate synthase; gene promoter; Rhodobacter capsulatus; transformant;
 KW MC9R/pRSFAC; ss.
 XX
 OS Escherichia coli.
 FH Key
 FT Location/Qualifiers
 CDS 380..877
 FT /*tag= a
 FT /label= ubiC
 FT 890..1782
 FT /*tag= b
 FT /label= ubiA
 FT 1868..1887
 FT /*tag= c
 FT /rpt_type= inverted
 FT 1892..1911
 FT /*tag= d
 FT /rpt_type= inverted
 XX
 PN JP08107789-A.
 XX
 XX
 PD 30-APR-1996.
 XX
 PF 13-OCT-1994; 94JP-00273071.
 XX
 PR 13-OCT-1994; 94JP-00273071.
 XX
 XX (ALPHA-) ALPHA SHOKUIN KK.
 PA
 XX
 DR WPI; 1996-262599/27.
 DR P-PSDB; AAR97746, AAR97747.
 XX
 XX Genes coding for ubiquinone biosynthetic enzymes - useful for prodn. of
 PT ubiquinone-10 by transformed photosynthetic bacteria.
 XX
 XX Claim 1; Fig 2; 11pp; Japanese.
 PS
 XX The present sequence comprises the ubiquinone biosynthetic enzyme genes
 CC ubiC and ubiA, which were cloned from the chromosomal DNA of E. coli
 CC (Kohara map phase DNA bank IF8 (634); Cell 50, 495-508 (1987)). A large

CC amt. of ubiquinone-10 can be produced by culturing photosynthetic
 CC bacteria transformed with the novel plasmid pRSFAC, which contains the
 CC ubiC and ubiA genes and the glutamate synthase gene promoter,
 CC specifically the Rhodobacter capsulatus transformant MC9R/pRSFAC
 XX
 SQ Sequence 2000 BP; 444 A; 462 C; 538 G; 556 T; 0 U; 0 Other;
 Alignment Scores: 2.98e-84 Length: 2000
 Pred. No.: 853.00 Matches: 165
 Score: 97.7% Conservative: 2
 Percent Similarity: 96.5% Mismatches: 4
 Best Local Similarity: 73.6% Indels: 0
 Query Match: 2 Gaps: 0
 DB: 2
 US-10-718-311-8 (1-227) x AAT29820 (1-2000)
 QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
 Db 362 TGTACCGGAGAGTTCGGCATGTACACCCCGGTTAACGCCAATGCTGCGCTGCGCTAT 421
 QY 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
 Db 422 TGTAAAGAGATCCCTGCCCTGGATCCGCAACTGCTGACTGGCTGTGCTGGAGGATTCC 481
 QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
 Db 482 ATGACAAAACGTTTTTGAACAGCAGGAGGAAAAACGGTAAGCGTGACGATGATGCCGAGGG 541
 QY 117 PheValGluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
 Db 542 TTGTGTCGAGCAGAAATGAAATCCCGAAGAACTGCGCGCTGCTGCCGAAAGAGTCTCGTTAC 601
 QY 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
 Db 602 TGGTTACGTGAATTTTGTATGTCCGATGGTGAACCGTGGCTTCCCGCTCGTACCGTC 661
 QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
 Db 662 GTTCCTGTCTCAACGTTAAGCGGGCGGAGCTGCCGCTTACAAAAATTTGGGTAACGCCG 721
 QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg 196
 Db 722 TTAGGACGCTATCTGTTTCATCATCATGACATTAACCCGGGACTTTATTGAGATAGCCGT 781
 QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
 Db 782 GATCGCGGCTGTGGGGCGGAGCTTCCCGCTCGGATTAAAGCGGTAACCGCTGTTGCTA 841
 QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
 Db 842 ACAGAACTGTTTTTACCGGCGTCACCGTTGTAC 874
 RESULT 14
 ADV91643
 ID ADV91643 standard; DNA; 1971 BP.
 XX
 AC ADV91643;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE C. parapsilosis pH8 1-H- E. coli CPL chimeric DNA.
 XX
 KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
 KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
 KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;
 KW genetically engineered microorganism; antioxidant; antimicrobial agent;
 KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cycostatic;
 KW neoplasm; pHBA 1-hydroxylase; pHBA; para-hydroxybenzoic acid;
 KW p-hydroxybenzoic acid; gene; ds.
 XX
 OS Candida parapsilosis.
 OS Escherichia coli.

OS Chimeric.
 FH Key Location/Qualifiers
 FT CDS 4..501 /*tag= a
 FT /product= "Escherichia coli CPL protein"
 FT /transl_except= (pos: 46..48, aa:Cys)
 FT 522..11961
 FT /*tag= b
 FT /product= "Candida parapsilosis pHEA 1-H protein"
 FT /transl_except= (pos: 540..542, aa:Trp)
 XX
 PN US2004261147-A1.
 XX
 XX
 PD 23-DEC-2004.
 XX
 XX 16-JUN-2003; 2003US-00462162.
 XX
 PF 16-JUN-2003; 2003US-00462162.
 XX
 PR (MEYE/) MEYER K.
 PA (VIIT/) VIITANEN P V.
 PA (FLIN/) FLINT D.
 XX
 XX Meyer K, Viitanen PV, Flint D;
 PI
 XX WPI; 2005-057232/06.
 DR P-PSDB; ADV91605, ADV91619.
 XX
 XX Producing hydroquinone glucoside in a green plant comprises growing a
 PT green plant having nucleic acid fragments and chorismate pyruvate lyase
 PT expression cassette.
 PT
 XX
 XX Example 7; SEQ ID NO 54; 70pp; English.
 PS
 XX The present invention relates to methods and materials to produce
 CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)
 CC in genetically modified green plants and microorganisms. The method
 CC relies upon transgenic plants or genetically modified microorganisms that
 CC produce increased levels of the initial substrate para-hydroxybenzoic
 CC acid (p-hydroxybenzoic acid; pHEA) in a biosynthetic pathway to produce
 CC arbutin. Plants and microbes can be genetically engineered to produce
 CC high levels of pHEA either by functional expression of the bacterial
 CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-
 CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.
 CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-
 CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The
 CC present sequence is Candida parapsilosis pHEA 1-hydroxylase (pHB 1-H) and
 CC Escherichia coli chorismate pyruvate lyase (CPL) chimeric DNA.
 XX
 SQ Sequence 1971 BP; 550 A; 383 C; 484 G; 554 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4,878-84 Length: 1971
 Score: 851.00 Matches: 165
 Percent Similarity: 99.4% Conservative: 0
 Best Local Similarity: 99.4% Mismatches: 1
 Query Match: 73.4% Indels: 0
 DB: 14 Gaps: 0
 US-10-718-311-8 (1-227) x ADV91643 (1-1971)

QY 122 GluIleProGluGluLeuProLeuProLysGluSerArgTyrTrpLeuArgGluIle 141
 DB |||||
 DB 181 GAATCCCGAAGAACTGCGCTGCGGAAAGAGTCTCGTTACTGGTTAGTGAATT 240
 QY 142 LeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrValValProValSerThr 161
 DB |||||
 DB 241 TTGTTATGTGCGGATGGTGAACCGTGGCTTGCCGCTCGTACCGTCTTCTGTGTCAAG 300
 QY 162 LeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyrLeu 181
 DB |||||
 DB 301 TTAAGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACGCCCTTAGGACCTATCTG 360
 QY 182 PheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeuTrp 201
 DB |||||
 DB 361 TTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGCCGTGATCGCGGGCTGTGG 420
 QY 202 GlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPheLeu 221
 DB |||||
 DB 421 GGGCGACGTTCCCGCTCGGATTAAAGCGTAAACCGCTGTTGCTAACAGAACTGTTTTTA 480
 QY 222 ProLaseProLeuTyr 227
 DB |||||
 DB 481 CCGCGTCCCGTGTGATC 498
 RESULT 15
 AAS90335
 ID AAS90335 standard; cDNA; 1207 BP.
 XX
 AC AAS90335;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #26139.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00849167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABG26148.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 26139; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (II) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 1207 BP; 316 A; 265 C; 288 G; 338 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.03e-80	Length:	1207
Score:	813.00	Matches:	164
Percent Similarity:	96.0%	Conservative:	2
Best Local Similarity:	94.8%	Mismatches:	5
Query Match:	70.1%	Indels:	2
DB:	5	Gaps:	0

US-10-718-311-8 (1-227) x AAS90335 (1-1207)

QY	57	CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr	76
DB	634	TGTAACGGAGAGTTCGGCATGTACACCCCGGTTAAACGCAACTGCGTGGCTAT	693
QY	77	CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer	96
DB	694	TGTAAGAGATCCCTGCCCTGGATCGCAACTGCTCGACTGGCTGTGTGGAGGATCC	753
QY	97	MetThrLysArgPheGlnGlnGlnGlyThrValSerValThrMetIleArgGluGly	116
DB	754	ATGACAAAACGTTTGAACAGCAGGAGAAACGTAAGCGTGACGATGATCCGCAAGGG	813
QY	117	PheValGluGlnAsnGluIleProGluGluLeuProLeuProLysGlnSerArgTyr	136
DB	814	TTTGTGAGCAGAAATGAAATCCCGAAGAACTGCGCGTCTGCCGAAAGAGTCTCGTTAC	873
QY	137	TrpLeuArgGluIleLeu-LeuCysAlaAspGlyGluProTrpLeuAla-GlyArgThrV	156
DB	874	TGGTTACGTGAATTTTGTGTATGTGCCGATGGCGNACCGCGGCTTGCCGGTCTGACCG	933
QY	156	alValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrP	176
DB	934	TCGTTCTCTGTCAACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGTAAACGC	993
QY	176	roLeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyA	196
DB	994	CGTTAGGACGCTATCTGTTACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCC	1053
QY	196	rgAspAlaGlyLeuTrpGlyArgArgSerArgLeuArgLeuSerGlyLysProLeuLeuL	216
DB	1054	GTGATGCCGGGCTGTGGGGCGGACGTTCCCGCTTCGATTAGCGGTAACCCGCTGTTC	1113
QY	216	euThrGluLeuPheLeuProAlaSerProLeuTyr	227
DB	1114	TAACAGAACTGTTTTACCGGGGTACACCGTTGTAC	1148

Search completed: July 11, 2006, 21:50:40
 Job time : 637.516 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
Run on: July 11, 2006, 19:46:18 ; Search time 4967.29 Seconds
(without alignments)
4383.493 Million cell updates/sec

Title: US-10-718-311-8
Perfect score: 1159
Sequence: 1 MASSWISSAAVATRSNVTOA.....RLSGRLLLTFLPASPLY 227

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=abs/ABSWEB spool/US10718311/runat.11072006.111520.5171/app.query.fasta.1
-DB=GenEmbl -QFWT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p
-USER=US10718311 @CEN 1 1 9663 @runat.11072006.111520.5171 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.env.*
2: gb.pat.*
3: gb.ph.*
4: gb.pl.*
5: gb.pr.*
6: gb.ro.*
7: gb.sts.*
8: gb.sy.*
9: gb.un.*
10: gb.vi.*
11: gb.ov.*
12: gb.htg.*
13: gb.in.*
14: gb.om.*
15: gb.ba.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Length	Description
1	1159	100.0	684 2 AR428736 Sequence
2	1159	100.0	684 2 AR455412 Sequence
3	1159	100.0	684 2 AX329368 Sequence

4	1159	100.0	684	2	AX382258	Sequence
5	856	73.9	854	2	AR144788	Sequence
6	856	73.9	1498	15	ECUBIA	M9268 Escherichia
7	856	73.9	2000	2	E11274	E11274 ubiC and ub
8	856	73.9	2000	15	ECUBICA	M93136 Escherichia
9	856	73.9	2000	15	ECUBIPLS	M93413 Escherichia
10	856	73.9	2034	15	ECUBIAC	X57434 E.coli ubiC
11	856	73.9	2348	15	ECUBI	X66619 E.coli gene
12	856	73.9	6641	2	AR144789	Sequence
13	856	73.9	110000	15	U00096	Continuation (43 o
14	856	73.9	110000	15	AP009048	Continuation (43 o
15	856	73.9	176195	15	ECUM89	U00006 E. coli chr
16	854	73.7	495	2	A47932	Sequence 1
17	854	73.7	495	2	AR428733	Sequence
18	854	73.7	495	2	AR455409	Sequence
19	854	73.7	495	2	AX329364	Sequence
20	854	73.7	495	2	AX382254	Sequence
21	852	73.5	110000	15	AE005174	Continuation (52 o
22	848	73.2	110000	15	BA000007	Continuation (51 o
23	848	73.2	110000	15	BA000007	Continuation (52 o
24	845	72.9	110000	15	AE005674	Continuation (44 o
25	845	72.9	110000	15	AE014073	Continuation (35 o
26	845	72.9	110000	15	CP000036	Continuation (42 o
27	843	72.7	1396	15	DO087228	DO087228 Escherich
28	841	72.6	110000	15	CP000034	Continuation (43 o
29	837	72.2	110000	15	CP000038	Continuation (45 o
30	835	72.0	110000	15	AE014075	Continuation (48 o
31	800	69.0	1918	2	E07849	E07849 gDNA encodi
32	725	62.6	23880	15	AE008898	AE008898 Salmonell
33	725	62.6	110000	15	CP000026	Continuation (42 o
34	725	62.6	110000	15	CP000026	Continuation (43 o
35	721	62.2	110000	15	AE014613	Continuation (44 o
36	721	62.2	110000	15	AE017220	Continuation (44 o
37	721	62.2	234050	15	AL627282	AL627282 Salmonell
38	673	58.1	645	2	AR384505	AR384505 Sequence
39	438.5	37.8	110000	15	BX936398	Continuation (5 of
40	437.5	37.7	10668	15	AE013658	AE013658 Versinia
41	437.5	37.7	214050	15	AJ414142	AJ414142 Versinia
42	437.5	37.7	290803	15	AE017128	AE017128 Versinia
43	424.5	36.6	110000	15	BX950851	Continuation (7 of
44	417	36.0	349107	15	BX571873	Continuation (7 of
45	408.5	35.2	110000	15	BX897699	Continuation (13 o

ALIGNMENTS

RESULT 1
AR428736
LOCUS AR428736 684 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 16 from patent US 6642036.
ACCESSION AR428736
VERSION AR428736.1 GI:40188466
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 684)
AUTHORS Flint,D., Meyer,K. and Viitanen,P.
TITLE Sinapoylglucose:malate sinapoyltransferase form malate conjugates from benzoic acid glucosides
JOURNAL Patent: US 6642036-A 16 04-NOV-2003;
E. I. Du Pont de Nemours and Company; Wilmington;
WOX;
FEATURES
source Location/Qualifiers
1..684
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN			
Alignment Scores:	Pred. No.:	Length:	
Score:	1159.00	684	
Percent Similarity:	100.0%	Matches:	227
		Conservative:	0

Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 2 Gaps: 0

US-10-718-311-8 (1-227) x AR428736 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGCCACACGAGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCCTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTACACCCCGGTTAACGCAACTGCTGCGCTATTTGTAAGAGATC 240
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGCGCTGGATCCGCAACTGCTGACCTGGCTGTGCTGGAGGATTCCATGACAAAACGT 300
Qy 101 PheGluGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db 301 TTTGAACAGCAGGGAACCGTAAAGCGTACGATGATCCGCGAAGGTTTGTGAGGAG 360
Qy 121 AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAAATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGTA 480
Qy 141 IleLeuLeuCysAlaAspGlyClnProTrpLeuAlaGlyArgThrValProValSer 160
Db 421 ATTTTGTATGTGCGCATGGTAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGGAGAGTGGCTGTACAAAATTGGGTAAACGCCGTTAGGACGCTAT 540
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTACATCATCGACATTAAACCGGACCTTTATTGAGATAGCGCGTATGCCGGCTG 600
Qy 201 TrpGlyArgSerArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT 660
Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGCGTCACCGTTGTAC 681

RESULT 2
LOCUS AR455412 684 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 7 from patent US 5683231.
ACCESSION AR455412
VERSION AR455412.1 GI:42689960
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 684)
AUTHORS Meyer, K., Viitanen, P.V. and Van Dyk, D.E.
TITLE High level production of p-hydroxybenzoic acid in green plants
JOURNAL Patent: US 5683231-A 7 27-JAN-2004;
E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES
Location/Qualifiers
1..684
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.12e-110 Length: 684
Pred. No.: 1159.00 Matches: 227
Score: 1159.00 Conservativity: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 100.0% Gaps: 0
DB: 2

US-10-718-311-8 (1-227) x AR455412 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTGTGCCACACGAGCAATGTTACACAAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTGGACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCCTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
Qy 61 TrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyrCysLysGluIle 80
Db 181 TGGCATATGTACACCCCGGTTAACGCAACTGCTGCGCTATTTGTAAGAGATC 240
Qy 81 ProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSerMetThrLysArg 100
Db 241 CTGCGCTGGATCCGCAACTGCTGACCTGGCTGTGCTGGAGGATTCCATGACAAAACGT 300
Qy 101 PheGluGlnGlyLysThrValSerValThrMetIleArgGluGlyPheValGluGln 120
Db 301 TTTGAACAGCAGGGAACCGTAAAGCGTACGATGATCCGCGAAGGTTTGTGAGGAG 360
Qy 121 AsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyrTrpLeuArgGlu 140
Db 361 AATGAAATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGTA 480
Qy 141 IleLeuLeuCysAlaAspGlyClnProTrpLeuAlaGlyArgThrValProValSer 160
Db 421 ATTTTGTATGTGCGCATGGTAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 161 ThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrProLeuGlyArgTyr 180
Db 481 ACGTTAAGCGGCGGAGAGTGGCTGTACAAAATTGGGTAAACGCCGTTAGGACGCTAT 540
Qy 181 LeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArgAspAlaGlyLeu 200
Db 541 CTGTTACATCATCGACATTAAACCGGACCTTTATTGAGATAGCGCGTATGCCGGCTG 600
Qy 201 TrpGlyArgSerArgLeuSerGlyLysProLeuLeuLeuThrGluLeuPhe 220
Db 601 TGGGGCGACGTTCCCGCTCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTT 660
Qy 221 LeuProAlaSerProLeuTyr 227
Db 661 TTACCGGCGTCACCGTTGTAC 681

RESULT 3
LOCUS AX329368 684 bp DNA linear PAT 08-JAN-2002
DEFINITION Sequence 7 from Patent WO0194607.
ACCESSION AX329368
VERSION AX329368.1 GI:18102367
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Meyer, K., van Dyk, D.E. and Viitanen, P.V.
TITLE High level production of p-hydroxybenzoic acid in green plants

RESULT 4			
AX382258			
LOCUS	AX382258	684 bp	DNA linear PAT 18-MAR-2002
DEFINITION	Sequence 16 from Patent WO0204653.		
ACCESSION	AX382258		


```

Pred. No.: 9 36e-79 Length: 1498
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 15 Gaps: 0

US-10-718-311-8 (1-227) x ECOUBIA (1-1498)

QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
   |||
   |||
Db 729 TGTAAACGGAGAGTTCGGCATGTACACCCCGCGTTAAACGCACTGCGTGGCGCTAT 788

QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
   |||
   |||
Db 789 TGTAAAGAGATCCCTGCCCTGGATCCGCACTCTCGACTGGCTGTGTCTGGAGGATTC 848

QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
   |||
   |||
Db 849 ATGACAAAACGTTTTGACACAGCAGGGAACCGTAAGCGTGACGATGATCCCGGAGGG 908

QY 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
   |||
   |||
Db 909 TTTGTGCGAGCAGATGAATCCCGAAGAACTCCGCTGCTGCGAAGAGATCTCGTTAC 968

QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
   |||
   |||
Db 969 TGGTTACGTGAATTTTGTATGTCGCGATGGTGAACCGTGGCTTCCCGCTCGTACCGTC 1028

QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
   |||
   |||
Db 1029 GTTCTGTGTCAACGTTAAGCGGCGGAGCTCGGCTTACAAAATTTGGTAAACCGCG 1088

QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
   |||
   |||
Db 1089 TTAGGACGCTATCTGTTACATCATCGACATTAACCGGGACITTTATGAGATAGCCGT 1148

QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
   |||
   |||
Db 1149 GATCGCGGGCTGTGGGGCGAGCTTCCGCTCGGATTAAGCGGTAACCGCTGTGTGCTA 1208

QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
   |||
   |||
Db 1209 ACAGAACTGTTTACCGGCGTACCCTGTGTAC 1241

RESULT 7
E11274
LOCUS E11274 2000 bp DNA linear PAT 29-SEP-1997
DEFINITION ubiC and ubiA gene.
ACCESSION E11274
VERSION E11274.1 GI:22024916
KEYWORDS JP 1996107789-A/1.
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 2000)
REFERENCE Matsuda,H., Kawamuki,M. and Nakagawa,T.
AUTHORS PRODUCTION OF UBIQUINONE-10.
JOURNAL Patent: JP 1996107789-A 1 30-APR-1996;
ALPHA- SHOKUHIN KK
OS Escherichia coli
PN JP 1996107789-A/1
PD 30-APR-1996
PF 13-OCT-1994 JP 1994273071
PI MATSUDA HIDEYUKI, KAWAMUKI MAKOTO, NAKAGAWA TSUYOSHI PC
(C12N15/09,A61K31/12,A61K31/12,C07H21/04,C12N1/21,C12P7/66, PC
(C12N1/21,
PC C12R1:01),(C12P7/66,C12R1:01);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH

```

```

FT source 1..2000
FT /organism='Escherichia coli'
FT CDS 380..877
FT /product='UbiC protein'
FT CDS 890..1762
FT /product='UbiA protein'.
FEATURES
    Location/Qualifiers
    1..2000
    /organism="Escherichia coli"
    /mol_type="genomic DNA"
    /db_xref="taxon:562"
ORIGIN
Alignment Scores: 1.33e-78 Length: 2000
Pred. No.: 856.00 Matches: 166
Score: 97.7% Conservative: 1
Percent Similarity: 97.1% Mismatches: 4
Best Local Similarity: 73.9% Indels: 0
Query Match: 2 Gaps: 0
DB:
US-10-718-311-8 (1-227) x E11274 (1-2000)
QY 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
   |||
   |||
Db 362 TGTAAACGGAGAGTTCGGCATGTACACCCCGCGTTAAACGCACTGCGTGGCGCTAT 421

QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
   |||
   |||
Db 422 TGTAAAGAGATCCCTGCCCTGGATCCGCACTGCTCGACTGGCTGTGTCTGGAGGATTC 481

QY 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
   |||
   |||
Db 482 ATGACAAAACGTTTTGACACAGCAGGGAACCGTAAGCGTGACGATGATCCCGGAGGG 541

QY 117 PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr 136
   |||
   |||
Db 542 TTTGTGCGACAGATGAATCCCGAAGAACTGCGCTGCTGCCGGAAGAGTCTCGTTAC 601

QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
   |||
   |||
Db 602 TGGTTACGTGAATTTTGTATGTCGCGATGGTGAACCGTGGCTTCCCGCTCGTACCGTC 661

QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
   |||
   |||
Db 662 GTTCTGTGTCAACGTTAAGCGGCGGAGCTGCGCTTACAAAATTTGGGTAAACCGCG 721

QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
   |||
   |||
Db 722 TTAGGACGCTATCTGTTACATCATCGACATTAACCGGGACITTTATGAGATAGCCGT 781

QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
   |||
   |||
Db 782 GATCGCGGGCTGTGGGGCGAGCTTCCGCTCGGATTAAGCGGTAACCGCTGTGTGCTA 841

QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
   |||
   |||
Db 842 ACAGAACTGTTTACCGGCGTACCCTGTGTAC 874

RESULT 8
ECOUBIA 2000 bp DNA linear BCT 26-APR-1993
LOCUS Escherichia coli chorismate lyase (ubiC), 4-hydroxybenzoate
DEFINITION octaprenyl transferase (ubiA) genes, complete cds, and
sn-glycerol-3-phosphate acyltransferase (plsB) genes, 3' end.
ACCESSION M93136
VERSION M93136.1 GI:148099
KEYWORDS 4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plsB
gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC
gene.
SOURCE Escherichia coli W3110
ORGANISM Escherichia coli W3110
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

```

REFERENCE 1 (bases 1 to 2000)
AUTHORS Lightner, V.A., Bell, R.M. and Modrich, P.
TITLE The DNA sequences encoding plbB and dgk loci of *Escherichia coli*
J. Biol. Chem. 258 (18), 10856-10861 (1983)
PUBMED 6309817

REFERENCE 2 (bases 1 to 2000)
AUTHORS Nichols, B.P. and Green, J.M.
TITLE Cloning and Sequencing of *Escherichia coli* ubiC and purification of chorismate lyase
Unpublished (1992)
JOURNAL Original source text: *Escherichia coli* (sub_strain W3110, strain K-12) (library: Kohare lambda-1F8) DNA.
COMMENT Location/Qualifiers
FEATURES
source
1..2000
/organism="Escherichia coli W3110"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="W3110"
/db_xref="taxon:316407"
/map="91.8 min"
/tissue_lib="Kohare lambda-1F8"
380..877
/gene="ubiC"
380..877
/gene="ubiC"
/function="enzymatic chorismate --> p-hydroxybenzoate + pyruvate"
/experiment="experimental evidence, no additional details recorded"
/citation=[2]
/codon_start=1
/transl_table=11
/product="chorismate lyase"
/protein_id="AAA24711.1"
/db_xref="GI:148100"
/translation="MSHPALTLQRLALRYCKEIPALDPQLLDMLLEDLSMTKRFQOQK TVSTMIREGFVEQNEIPEELPLLPKESRYLWREILLCADGEPWLAGRTVVPVSTLSG PELALQKLGKPLGRLYFTSITLDRDFEIGRDAGLWGRSRLSGKPLLLTLFLFPL ASPLV"
890..1762
/gene="ubiA"
890..1762
/gene="ubiA"
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation=[2]
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA24712.1"
/db_xref="GI:148101"
/translation="MNSLTQNKLLAFHRLMRTDKPIGALLLWPTLWALWVATPGVP QLMVIAFVAGVWLRAGCVVNDYADRFKDFGHVKRTANRELPSPGAVTERARALFV LVLIFLVLTLNTWITLLISAALAWVYPMKRYTHLPQVILGAAGFSGIPMAFAA VSESVPLSCWLMFLANIWAVYDTQYAMVDRDDVKIGIKSTAILFGYDKLLIGIL QIGVLALMAIGELNGLGWYWSILVAGALFVYQOKLIANREACPKAFMNNYVG LVLFLGLWSTWHF"
complement(1917..2000)
/gene="plbB"
complement(1917..2000)
/gene="plbB"
/EC_number="2.3.1.15"
/function="lysophosphatidic acid synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation=[1]
/codon_start=1
/transl_table=11
/product="sn-glycerol-3-phosphate acyltransferase"
/protein_id="AAA24713.1"

Enterobacteriaceae: *Escherichia*.
1 (bases 1 to 2000)
Lightner, V.A., Bell, R.M. and Modrich, P.
The DNA sequences encoding plbB and dgk loci of *Escherichia coli*
J. Biol. Chem. 258 (18), 10856-10861 (1983)
PUBMED 6309817

2 (bases 1 to 2000)
Nichols, B.P. and Green, J.M.
Cloning and Sequencing of *Escherichia coli* ubiC and purification of chorismate lyase
Unpublished (1992)
JOURNAL Original source text: *Escherichia coli* (sub_strain W3110, strain K-12) (library: Kohare lambda-1F8) DNA.
COMMENT Location/Qualifiers
FEATURES
source
1..2000
/organism="Escherichia coli W3110"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="W3110"
/db_xref="taxon:316407"
/map="91.8 min"
/tissue_lib="Kohare lambda-1F8"
380..877
/gene="ubiC"
380..877
/gene="ubiC"
/function="enzymatic chorismate --> p-hydroxybenzoate + pyruvate"
/experiment="experimental evidence, no additional details recorded"
/citation=[2]
/codon_start=1
/transl_table=11
/product="chorismate lyase"
/protein_id="AAA24711.1"
/db_xref="GI:148100"
/translation="MSHPALTLQRLALRYCKEIPALDPQLLDMLLEDLSMTKRFQOQK TVSTMIREGFVEQNEIPEELPLLPKESRYLWREILLCADGEPWLAGRTVVPVSTLSG PELALQKLGKPLGRLYFTSITLDRDFEIGRDAGLWGRSRLSGKPLLLTLFLFPL ASPLV"
890..1762
/gene="ubiA"
890..1762
/gene="ubiA"
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation=[2]
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA24712.1"
/db_xref="GI:148101"
/translation="MNSLTQNKLLAFHRLMRTDKPIGALLLWPTLWALWVATPGVP QLMVIAFVAGVWLRAGCVVNDYADRFKDFGHVKRTANRELPSPGAVTERARALFV LVLIFLVLTLNTWITLLISAALAWVYPMKRYTHLPQVILGAAGFSGIPMAFAA VSESVPLSCWLMFLANIWAVYDTQYAMVDRDDVKIGIKSTAILFGYDKLLIGIL QIGVLALMAIGELNGLGWYWSILVAGALFVYQOKLIANREACPKAFMNNYVG LVLFLGLWSTWHF"
complement(1917..2000)
/gene="plbB"
complement(1917..2000)
/gene="plbB"
/EC_number="2.3.1.15"
/function="lysophosphatidic acid synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation=[1]
/codon_start=1
/transl_table=11
/product="sn-glycerol-3-phosphate acyltransferase"
/protein_id="AAA24713.1"

/db_xref="GI:148102"
/translation="MKVYQLLAELITSDVRLTIESATOGEG"

ORIGIN
Alignment Scores:
Pred. No.: 1.33e-78 Length: 2000
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 15 Gaps: 0

US-10-718-311-8 (1-227) x ECOUBICA (1-2000)

Qy 57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
Db 362 TGTAAAGGAGAGTTCGGCATGTACACCCCGGTTAACGCAACTCGTGGCTGGCTAT 421

Qy 77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 422 TGTAAAGAGATCCCTCGCTGATCCGCACTGCTCGACTGGCTGGCTGGAGATTCC 481

Qy 97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db 482 ATGACAAACGTTTGTAAACACGACGAGGAAAAACGGTAAGCGTGACGATATCCGCAAGGG 541

Qy 117 PheValGluGlnAsnGluIleProGluGluLeuProGluLeuProLysGluSerArgTyr 136
Db 542 TTTGTGAGCAGATGAATCCCGNAGAACTCGCGTCTGCCGAAAGAGTCTGTTAC 601

Qy 137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 602 TGGTTACGTGAATTTTGTATGTCCGATGGTGAACCGTGGCTGGCTGGCTACCGTC 661

Qy 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 662 GTTCCTGTGTCAACGTTAAAGCGGCGGAGTGGCGTTACAAAAATTGGGTAAACGCGG 721

Qy 177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArg 196
Db 722 TTAGACGCTATCTGTTACATCATCGACATTACCCGGGACTTTATTGAGATAGCCGT 781

Qy 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
Db 782 GATCGCGGCTGTGGGGCGACGTTCCCGCTCGGATTAAAGCGGTAAACCGCTGTGCTA 841

Qy 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 842 ACAGAACTGTTTTTACC GGCGGTACCGGTGTAC 874

RESULT 9
ECOUBPLS 2000 bp DNA linear BCT 26-APR-1993
LOCUS *Escherichia coli* 4-hydroxybenzoate octaprenyl transferase (ubiA)
DEFINITION gene complete cds, chorismate lyase (ubiC) gene complete cds, sn-glycerol-3-phosphate acyltransferase (plbB) gene, 3' end.
M93413
ACCESSION M93413.1 GI:148106
VERSION 4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plbB gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC gene.
KEYWORDS *Escherichia coli* W3110
SOURCE *Escherichia coli* W3110
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; *Escherichia*.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Lightner, V.A., Bell, R.M. and Modrich, P.
TITLE The DNA sequences encoding plbB and dgk loci of *Escherichia coli*
J. Biol. Chem. 258 (18), 10856-10861 (1983)
PUBMED 6309817
REFERENCE 2 (bases 1 to 2000)
AUTHORS Nichols, B.P. and Green, J.M.
TITLE Cloning and sequencing of *Escherichia coli* ubiC and purification of

chorismate lyase
J. Bacteriol. 174 (16), 5309-5316 (1992)
1644758
Original source text: Escherichia coli (sub_strain W3110, strain K-12) (library: Kohara lambda) DNA.
Location/Qualifiers
1..2000
/organism="Escherichia coli W3110"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="W3110"
/db_xref="taxon:316407"
/map="91.8 min"
/tissue_lib="Kohara lambda"
380..877
/gene="ubiC"
380..877
/gene="ubiC"
/function="enzymatic chorismate --> p-hydroxybenzoate + pyruvate"
/experiment="experimental evidence, no additional details recorded"
/note="N-terminal amino acid sequence of chorismate lyase -NH2-SHPALTLRALRYCEIPAL-"
/citation=[2]
/codon_start=1
/transl_table=11
/product="chorismate lyase"
/protein_id="AAA24716.1"
/db_xref="GI:148107"
/translation="MSHPALTLRALRYCEIPALDPQLLDWLLLEDSTKRFQOQK TVSVTMIRGFVEQNEIPEELPLPKESRYWUREILLCADGEPFAGRTVVPVSLSG PELALQKLKPTPLGRYLTSTLTDFIEIGRDAGLWGRSRLSLGKPLLELFLP ASPLY"
890..1762
/gene="ubiA"
890..1762
/gene="ubiA"
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation=[2]
/number=2
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA24717.1"
/db_xref="GI:148108"
/translation="MEWSLTQNKLLAFHRLMTRDKPIGALLLLNPTLWALWVATPGVP QWLILAVFVAGVLMRACGVYNDYADRFKDFGDKVKTANRPLPSGAVTEKEARALFV LVLISFLVLTNTMTILSIAALAWYPMKRYTHLPQVVLGAFAAFIPMAFAA VSESPLSCWLMFLANLWAVDTQYAMVDRDDDKVIGIKSTAILFGYDKLIIGIL QIGVALAMAIIGELNGLGWYWSILVAGALFYQOKLIANREERACFKAFNNVVG LVLFLGLAMSYWHF"
complement(1917..2000)
/gene="plsB"
complement(1917..2000)
/gene="plsB"
/EC_number="2.3.1.15"
/function="lysophosphatidic acid synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation=[1]
/transl_table=11
/codon_start=1
/product="sn-glycerol-3-phosphate acyltransferase"
/protein_id="AAA24718.1"
/db_xref="GI:148109"
/translation="WKVYQLLAELITSDVRLTIESATQEGE"
ORIGIN
Alignment Scores:

Pred. No.: 1,338-78 Length: 2000
Score: 856.00 Matches: 166
Percent Similarity: 97.7% Conservative: 1
Best Local Similarity: 97.1% Mismatches: 4
Query Match: 73.9% Indels: 0
DB: 15 Gaps: 0
US-10-718-311-8 (1-227) x ECOUBIPLS (1-2000)
QY 57 CysMetGinValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr 76
Db 362 TGTAACGAGAGAGTTCGGCATGTACACCCCGGCTTAAGCAACTGCGTGCCTGCCTAT 421
QY 77 CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96
Db 422 TGTAAGAGATCCCTGCGCTGGATCCCAACTGCTCGACTGCTGTGTGGAGGATTC 481
QY 97 MetThrLysAspPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116
Db 482 ATGACAAAACGTTTTGAACAGCAGGAGGAAAACGGTAAGGTGACGATGATCCGGAAGG 541
QY 117 PheValGluGlnAsnGluLeuProGluLeuProLeuLeuProLysGluSerArgTyr 136
Db 542 TTTGTCGAGCAGCAATGAAATCCCGAAGAACTGCCGCTGCTGCCAAAGAGTCTCGTTAC 601
QY 137 TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal 156
Db 602 TGGTTACGTGAAATTTTGTATGTCGATGCGTGAACCGTGGCTTGCCTGATACCGTC 661
QY 157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176
Db 662 GTTCCTGTGTAACGTTAAGCGCGCGGAGCTGGGCTTACAAAATTGGGTAAACGCCG 721
QY 177 LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluLeuGlyArg 196
Db 722 TTAGGACGCTATCTGTTACATCATCATGACATTAACCCGGGACTTTATTGAGATAGCCGT 781
QY 197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216
Db 782 GATGCGCGGCTGTGGGGCGAGCTTCCCGCTGCGGATTAAGCGGTAACCGCTGTGTGCTA 841
QY 217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227
Db 842 ACAGAACTGTTTTTACCGCGCTACCGTTGTATAC 874
RESULT 10
ECUBIAC
LOCUS
DEFINITION
E.coli ubiC and ubiA genes for chorismate lyase and 4-hydroxybenzoate octaprenyltransferase.
X57434
VERSION
X57434.1 GI:43233
KEYWORDS
4-hydroxybenzoate octaprenyl transferase; 4-hydroxybenzoate synthetase; chorismate lyase.
SOURCE
Escherichia coli
ORGANISM
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE
1
AUTHORS
Nishimura, K., Nakahigashi, K. and Inokuchi, H.
TITLE
Location of the ubiA gene on the physical map of Escherichia coli
J. Bacteriol. 174 (17), 5762 (1992)
PUBMED
1512213
REFERENCE
2 (bases 1 to 2034)
AUTHORS
Nishimura, K.
TITLE
Direct Submission
JOURNAL
Submitted (28-JAN-1991) K. Nishimura, Faculty of Sciences, Kyoto University, Sakyo-ku, Kyoto 606, Japan
LOCATION/Qualifiers
1..2034
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12 W3110"

-35_signal -10_signal gene	/db_xref="taxon:562"	
	/map="91.7 min"	
CDS	/clone="634"	
	/clone_lib="Kohara library"	
CDS	137..142	
	159..164	
CDS	291..899	
	291..899	
CDS	/gene="ubiC"	
	/gene="ubiC"	
CDS	/codon_start=1	
	/transl_table=11	
CDS	/product="4-hydroxybenzoate synthetase"	
	/protein_id="CAA40681.1"	
CDS	/db_xref="GI:43234"	
	/db_xref="GOA:P26602"	
CDS	/db_xref="UniProtKB/Swiss-Prot:P26602"	
	/translation="MRLRFCCVDHLCFTSPVNTFLRYNAFTLCNGCFQMSHPALT QLRALRYCKEIPALDPQLLDLWLLDSMTKRFQOGKTVSMIREGFVEQNEIPEEL PLLPKESRYWLREILLCADGEPWLAGRTVPVPVSTLSPGPELALQKLGKTLPLGRYLTSS TLTRDFEIGRDAGLWGRRLRLSGKPLLLTELFLPASPLY"	
CDS	912..1784	
	/gene="ubiA"	
CDS	912..1784	
	/gene="ubiA"	
CDS	/codon_start=1	
	/transl_table=11	
CDS	/product="4-hydroxybenzoate octaprenyltransferase"	
	/protein_id="CAA40682.1"	
CDS	/db_xref="GI:43235"	
	/db_xref="GOA:P26601"	
CDS	/db_xref="UniProtKB/Swiss-Prot:P26601"	
	/translation="MWSLTQNKLIADFRLMRTDKPIGALLLWPTLMALWVATPGVP QLWTLAVAGVWLMRAAGCVNDYADRKFDGKHVRTANRPLPSGAVTEKARALFV LVLSFLLPLNTWLTLLSIAALAWVPFMKRYTHLPQVILGAAGFWSIPWFAA VSEVPLSCWLMFLANILWAVDTQYAMVDRDDDKIGIKSTAILFGQYDKLLIGIL QIGVLAMAIIGELNGWGYWSILVAGALFVYQQKLIANREREAFCFAPMNNYVG LVLFGLAMSWHF"	
CDS	ORIGIN	
	Alignment Scores:	
CDS	Pred. No.: 1,368-78 Length: 2034	
	Score: 856.00 Matches: 166	
CDS	Percent Similarity: 97.7% Conservative: 1	
	Best Local Similarity: 97.1% Mismatches: 4	
CDS	Query Match: 73.9% Indels: 0	
	DB: 15 Gaps: 0	
CDS	US-10-718-311-8 (1-227) x ECUBIAC (1-2034)	
	57 CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuAlaLeuArgTyr 76	
CDS	384 TGTAACGGAGAGTTCGGCATGTACACCCCGGTTAACGCAACTGCGTGCCTAT 443	
	77 CysLysGluIleProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuGluAspSer 96	
CDS	444 TGTAAGAGATCCCTCGCCTGATCCGCAACTGCTCGACTGCTGCTTCTGGAGATTCC 503	
	97 MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly 116	
CDS	504 ATGACAAACGTTTTTGAACAGCAGGAGAAAACGGTAAACGGTACGATGATCGCGAAGGG 563	
	117 PheValGluGlnAsnGluIleProGluGluLeuProLeuLeuProLysGluSerArgTyr 136	
CDS	564 TTTGTGCGAGAGATGAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTGTTAC 623	
	137 TrpLeuArgGluIleLeuLeuCysAlaAspGlyGluProTrpTrpLeuAlaGlyArgThrVal 156	
CDS	624 TGGTTACGTGAATTTTGTATGTCCGATGGTGAACCGTGGCTTGGCCGGTCGTACCGTC 683	
	157 ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro 176	
CDS	684 GTTCCTGTGTACACGTTAAGCGGGCCGAGCTGGCGTTACAAAATTTGGGTAAACGCGC 743	
QY	177 LeuGlyArgTyrLeuPheThrSerThrLeuThrArgAspPheIleGluIleGlyArg 196	
	744 TTAGAGCGTATCTGTTTCACATCATCGACATTAACCGGGAGCTTTATTGAGATAGCCGT 803	
QY	197 AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216	
	804 GATCCGGGCTGTGGGGCGACGTTCCCGCTCGCATTAAGCGGTAAACCGCTGTTGCTA 863	
QY	217 ThrGluLeuPheLeuProAlaSerProLeuTyr 227	
	864 ACAGACTGTTTTTACCGCGCTCACCGTTGTAC 896	
QY	ECUBI	
	2348 bp DNA linear BCT 09-SEP-2004	
QY	ECUBI	
	E.coli genes ubiC and ubiA.	
QY	X66619.1	
	X66619.1 GI:43230	
QY	4-hydroxybenzoate-octaprenyltransferase; chorismate-pyruvate lyase;	
	ubiA gene; ubiC gene.	
QY	Escherichia coli	
	Escherichia coli	
QY	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
	Enterobacteriaceae; Escherichia.	
QY	1 (bases 1 to 2348)	
	Siebert, M., Bechthold, A., Melzer, M., May, U., Berger, U.,	
QY	Schroder, G., Schroder, J., Severin, K. and Heide, L.	
	Ubiquinone biosynthesis. Cloning of the genes coding for chorismate	
QY	pyruvate-lyase and 4-hydroxybenzoate octaprenyl transferase from	
	Escherichia coli	
QY	FEBS Lett. 307 (3), 347-350 (1992)	
	1644192	
QY	2 (bases 1 to 2348)	
	Heide, L.	
QY	Direct Submission	
	Submitted (04-JUN-1992) L. Heide, Inst of Pharmaceutical Biology,	
QY	Schaezlestrasse 1, 7800 Freiburg, FRG	
	On Sep 24, 2004 this sequence version replaced gi:253104.	
QY	See also K00127, M93136 & M93413.	
	Location/Qualifiers	
QY	1..2348	
	/organism="Escherichia coli"	
QY	/mol_type="genomic DNA"	
	/strain="MC4100"	
QY	/db_xref="taxon:562"	
	/map="91.8 min"	
QY	/clone="PALMUI"	
	/clone_lib="Sau3A partial digested DNA in pbluescript"	
QY	558..563	
	570..1067	
QY	/gene="ubiC"	
	570..1067	
QY	/gene="ubiC"	
	/codon_start=1	
QY	/transl_table=11	
	/product="chorismate-pyruvate lyase"	
QY	/protein_id="CAA47181.1"	
	/db_xref="GI:43231"	
QY	/db_xref="GOA:P26602"	
	/db_xref="UniProtKB/Swiss-Prot:P26602"	
QY	/translation="MSPFALQRLALRYCKEIPALDPQLLDLWLLDSMTKRFQOGK TVSTMIREGFVEQNEIPEELPLLPKESRYWLREILLCADGEPWLAGRTVPVPVSTLSG PELALQKLGKTLPLGRYLTSSLTTRDFEIGRDAGLWGRRLRLSGKPLLLTELFLP ASPLY"	
	1068..1073	
QY	1080..1952	
	/gene="ubiA"	
QY	1080..1952	
	/gene="ubiA"	
QY	/codon_start=1	
	/transl_table=11	
QY	/product="4-hydroxybenzoate-octaprenyl transferase"	

DB:	15	Caps:	0	source
US-10-718-311-8 (1-227) x AP009048_42 (1-110000)				
QY	57	CysMetGlnValTrpHisMetSerHisProAlaLeuThrGlnLeuArgAlaLeuArgTyr	76	/organism="Escherichia coli"
Db	56078	TGTAACGGAGAGTTCGGCATGTACACCCCGGTAAACGCACTGCGTGGCTGCCTAT	56137	/mol_type="unassigned DNA"
QY	77	CysLysGluLeuProAlaLeuAspProGlnLeuLeuAspTrpLeuLeuLeuGluAspSer	96	/strain="K-12"
Db	56138	TGTAAGAGATCCCTCCCTGGATCGCGCACTGCTGACTGGCTGTGTCTGGAGGATCC	56197	/sub_strain="MG1655"
QY	97	MetThrLysArgPheGluGlnGlnGlyLysThrValSerValThrMetIleArgGluGly	116	/db_xref="taxon:562"
Db	56198	ATGACAAAACGTTTTGAACAGACGAGGAAACCGTAAGCGTGACGATGATCCGGAAGGG	56257	/note="this sequence comprises the following lambda clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203, EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262, EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27; M13mp19 or Janus vectors were used for subcloning"
QY	117	PheValGluGlnAsnGluLeuProGluGluLeuProLeuLeuProLysGluSerArgTyr	136	<1. 1737
Db	56258	TTTGTCGAGCAGAAATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTCGTTAC	56317	/note="corresponds to M21516; ECOKATGA(1225. .2805)"
QY	137	TrpLeuArgGluLeuLeuLeuCysAlaAspGlyGluProTrpLeuAlaGlyArgThrVal	156	1..1257
Db	56318	TGGTACGTAATTTTGTATGTCGCGATGTTGAACCGTGGCTTCCCGTCTGATCCGTC	56377	/gene="katG"
QY	157	ValProValSerThrLeuSerGlyProGluLeuAlaLeuGlnLysLeuGlyLysThrPro	176	/gene="katG"
Db	56378	GTTCTGTGTCAACGTTAAGCGCGCGGAGCTGGCGTTACAAAATTGGGTAAACGCCG	56437	/EC_number="1.11.1.6"
QY	177	LeuGlyArgTyrLeuPheThrSerSerThrLeuThrArgAspPheIleGluIleGlyArg	196	/standard_name="catalase HPI; catalase-peroxide; catalase"
Db	56438	TTAGACGCTACTCTGTTACATCATCATGACATTAACCCGGGACTTATTGAGATAGCCGT	56497	/note="CG_Site No. 14983"
QY	197	AspAlaGlyLeuTrpGlyArgSerArgLeuArgLeuSerGlyLysProLeuLeuLeu 216	56557	/product="catalase hydroperoxidase I"
Db	56498	GATCCCGGCTGTGGGGCGAGCTTCCCGCTGCGATTAAAGCGTAAACCGCTGTGCTA	56557	/protein_id="AAC43048.1"
QY	217	ThrGluLeuPheLeuProAlaSerProLeuTyr 227		/db_xref="GI:396289"
Db	56558	ACAGAACTGTTTACCGCGGTCACCGTGTGATC 56590		/translation="AITSGLVVTQTPTQWSNYFENLFKYEWTQTRSPAGAIQFEA VTHPSIIDPDPFSKKRPMTLVTLTLRFDPEFKISRFNLDPOAFNEAFARXWFK LTHRDMGPKSRVIGPEVPKEDLIWQDPLPQPIYNTEQDIIDLKFAIDSGLSVSLV SVWASASTFRGGDKRGAGLARLALPORDVNAAVRALPVLEKIOKESGKSLA DIIVLVGVGVKEAASAGLSIHPFAPGRVARDODTDIEMFELLEPIADGPRNVRA RLDVSTTSLDIDKAQQLTLTAPENTALVGGMRVLGANGFNEDGSKNGVFTDRVGLVSNDF RLNDMRYEAKDESKELEFGRDRETEGVEKFTASRADLVFGSNVLRVAEVIASS DAHEKFDKDFVAAMVKVMNLDKFDLL"
RESULT 15				
ECOW89				
LOCUS	ECOW89 176195 bp DNA linear BCT 17-DEC-1993			
DEFINITION	E. coli chromosomal region from 89.2 to 92.8 minutes.			
ACCESSION	U00006			
VERSION	U00006.1 GI:409785			
KEYWORDS	Escherichia coli			
SOURCE	Escherichia coli			
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.			
REFERENCE	Blattner, F.R., Burland, V., Plunkett, G. III, Sofia, H.J. and Daniels, D.L.			
AUTHORS	Analysis of the Escherichia coli genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes			
TITLE	Nucleic Acids Res. 21 (23), 5408-5417 (1993)			
JOURNAL	8265357			
PUBMED	2 (bases 1 to 176195)			
REFERENCE	Blattner, F.R.			
AUTHORS	Direct Submission			
TITLE	Submitted (03-SEP-1993) 608-263-7459			
JOURNAL	On Oct 29, 1993 this sequence version replaced gi:396288.			
PUBMED	This sequence was determined as part of the E. coli Genome Project (Frederick R. Blattner, director) at the University of Wisconsin-Madison. Supported by award HG00301 from the NIH Human Genome Project. The entire sequence was independently determined from E. coli MG1655; overlaps and conflicts with other sequence determinations are annotated. The start of this entry overlaps the end of the entry ECOW87 (119201) by 93 bp. NOTE: An update was submitted on 25-OCT-1993, reflecting a correction to the heme sequence.			
REFERENCE	Location/Qualifiers			
AUTHORS	EYPQKWLKRPDSNIDHRRVNPLETWFSRHDKTRPTSKNPSDYQAGDIVSWRLDGLA			

THIS PAGE BLANK (USPTO)

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 11, 2006, 19:21:37 ; Search time 137 Seconds
(without alignments)
6491.174 Million cell updates/sec
Title: US-10-718-311-7
Perfect score: 684
Sequence: 1 atgggtctctctgtctatttc.....cggcgctaccgtgtactaa 684
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 844069 seqs, 650066433 residues
Total number of hits satisfying chosen parameters: 1688138
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : Published Applications NA New.*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq.*
9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	21.1	762	US-11-247-587-125	Sequence 125, App
2	142.4	20.8	781	US-11-247-587-124	Sequence 124, App
3	124.6	18.2	698	US-11-247-587-128	Sequence 128, App
4	121.4	17.7	683	US-11-247-587-126	Sequence 126, App
5	119.8	17.5	700	US-11-247-587-127	Sequence 127, App
6	109.8	16.1	727	US-11-247-587-129	Sequence 129, App
7	85.2	12.5	718	US-10-953-349-10521	Sequence 10521, A
8	83.6	12.2	583	US-11-218-305-17631	Sequence 17631, A
9	83.6	12.2	658	US-11-218-305-17633	Sequence 17633, A
10	82	12.0	743	US-10-953-349-8894	Sequence 8894, App
11	80.4	11.8	640	US-11-218-305-17634	Sequence 17634, A
12	78.4	11.5	1023	US-11-216-545-4174	Sequence 4174, App
13	76.8	11.2	841	US-11-216-545-4173	Sequence 4173, App
14	76.2	11.1	1237	US-11-216-545-8381	Sequence 8381, App
15	74.4	10.9	799	US-10-953-349-9569	Sequence 9569, App
16	74.4	10.9	828	US-11-218-305-17632	Sequence 17632, A
17	46.2	6.8	627	US-11-216-545-6801	Sequence 6801, App
18	46.2	6.8	862	US-11-266-748A-36894	Sequence 36894, App
19	46.2	6.8	882	US-11-266-748A-388679	Sequence 388679, App
20	46.2	6.8	882	US-11-266-748A-450273	Sequence 450273, App
21	46.2	6.8	947	US-10-449-902-24132	Sequence 24132, A
22	39.8	5.8	845	US-10-449-902-19918	Sequence 19918, A
23	39.8	5.8	1999	US-10-449-902-17846	Sequence 17846, A
24	39.6	5.8	830	US-10-449-902-4174	Sequence 4174, App

25	38.2	5.6	2409	6	US-10-449-902-20231	Sequence 20231, A
26	37.8	5.5	899	6	US-10-449-902-18147	Sequence 18147, A
27	35.6	5.2	792	8	US-11-217-529-6134	Sequence 6134, App
c 28	35.6	5.2	1735	8	US-11-216-545-7862	Sequence 7862, App
c 29	33	4.8	1459	8	US-11-266-748A-381174	Sequence 381174, App
c 30	33	4.8	1459	8	US-11-266-748A-430930	Sequence 430930, App
c 31	32.8	4.8	1000	8	US-11-266-748A-222987	Sequence 222987, App
c 32	32.8	4.8	1000	8	US-11-266-748A-289901	Sequence 289901, App
c 33	32.8	4.8	1000	8	US-11-266-748A-292596	Sequence 292596, App
c 34	32.8	4.8	1000	8	US-11-266-748A-341330	Sequence 341330, App
c 35	32.8	4.8	1000	8	US-11-266-748A-344025	Sequence 344025, App
c 36	32.8	4.8	1000	8	US-11-266-748A-390405	Sequence 390405, App
c 37	32.8	4.8	1000	8	US-11-266-748A-401025	Sequence 401025, App
c 38	32.8	4.8	1000	8	US-11-266-748A-404362	Sequence 404362, App
c 39	32.8	4.8	1000	8	US-11-266-748A-404363	Sequence 404363, App
c 40	32.8	4.8	1000	8	US-11-266-748A-472071	Sequence 472071, App
c 41	32.8	4.8	1000	8	US-11-266-748A-475408	Sequence 475408, App
c 42	32.8	4.8	1000	8	US-11-266-748A-475409	Sequence 475409, App
c 43	32.8	4.8	1000	8	US-11-266-748A-481123	Sequence 481123, App
c 44	32.8	4.8	4378	8	US-11-266-748A-23487	Sequence 23487, A
c 45	32.8	4.8	4612	8	US-11-266-748A-234708	Sequence 24708, A

ALIGNMENTS

RESULT 1

US-11-247-587-125
; Sequence 125, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 125
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-125

Query Match	21.1%	Score 144;	DB 8;	Length 762;
Best Local Similarity	86.4%	Pred. No. 8.7e-39;		
Matches 159;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
QY	1	ATGGCTTCTCTGTCTCAATTTCTTTCAGCAGCTGTGTTCACACGACGCAATGTTCACAAAGCT	60	
DB	27	ATGGCTTCTCTGTCTCAATTTCTTTCAGCAGCTGTGTTCACACGACGCAATGTTCACAAAGCT	86	
QY	61	AGATGGTTGACCTTTCATCTGGTCTCAATTTTCAGCAGCTGTGTTCACACGCAATGTTCACAAAGCT	120	
DB	87	AAATGGTTGACCTTTCATCTGGTCTCAATTTTCAGCAGCTGTGTTCACACGCAATGTTCACAAAGCT	146	
QY	121	CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTTCAGTTCATGTCAGGCTG	180	
DB	147	CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTTCAGTTCATGTCAGGCTG	206	
QY	181	TGCG 184		
DB	207	TGCG 210		

```
RESULT 2
US-11-247-587-124
; Sequence 124, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 124
; LENGTH: 781
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-124
```

```
Query Match 20.8%; Score 142.4; DB 8; Length 781;
Best Local Similarity 85.9%; Pred. No. 3.1e-38;
Matches 158; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCTTTGCGACACGAGCAATGTTACACAAGCT 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 ATGGCTTCTCTGAGTTCTTCTTCAGCAGAGTGTGCGACCGCAGCAATGTTGCTCAAGCT 98

Qy 61 AGCATGGTTGACCTTTCAATCTTCTCAGCCACTTTCCCTGTTTACAAGAAG 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 AACATGGTTGACCTTTACAGCTTTAAGTCTGCTCATTCCTGCTTCAAGAAG 158

Qy 121 CAAACCTTGACATCACTTCCATCTAGCAATGTTGGAAGAGTTAGTGCATGCAAGTG 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 CAAACCTTGACATCACTTCCATCTGCGACGCGGAGAGTGCATGCAAGTG 218

Qy 181 TGGC 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
219 TGGC 222
```

```
RESULT 3
US-11-247-587-128
; Sequence 128, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
```

```
; SEQ ID NO 128
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-128

Query Match 18.2%; Score 124.6; DB 8; Length 698;
Best Local Similarity 82.9%; Pred. No. 3.7e-32;
Matches 155; Conservative 0; Mismatches 29; Indels 3; Gaps 1;

Qy 1 ATGGCTTCTCTGTCATTTCTTCAGCAGCT---GTTGCCACACGAGCAATGTTACACAA 57
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20 ATGGCTTCTCTGAGTTCTCTCAGCTGCTGCGTGGCACCGCGGCTAATGCTGCTCAA 79

Qy 58 GCTAGCATGGTTGCACCTTTCACTGGTCTCAAACTCTTCAGCCACTTTCCCTGTTTACAAAG 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
80 GCCAGCATGGTTGCACCTTTCACTGGCCTCAAACTCTCTCTCTCCCTGTTTACCAGA 139

Qy 118 AAGCAAAACCTTGACATCACTTCCATCTAGCAATGTTGGAAGAGTTAGTGCATGCGAG 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
140 AAACAAACCTTGACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 199

Qy 178 GTGTGGC 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200 GTGTGGC 206
```

```
RESULT 4
US-11-247-587-126
; Sequence 126, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 126
; LENGTH: 683
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-126
```

```
Query Match 17.7%; Score 121.4; DB 8; Length 683;
Best Local Similarity 81.8%; Pred. No. 4.6e-31;
Matches 153; Conservative 0; Mismatches 31; Indels 3; Gaps 1;

Qy 1 ATGGCTTCTCTGTCATTTCTTCAGCA---GCTTTGCCACACGAGCAATGTTACACAA 57
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
37 ATGGCTTCTCTCAGTTATGTCCTCAGCAGCTGCTTTGCGACCGCGGCAATGCTGCTCAA 96

Qy 58 GCTAGCATGGTTGCACCTTTCACTGGTCTCAAACTCTTCAGCCACTTTCCCTGTTTACAAAG 117
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
97 GCCAACATGGTTGACCCCTTCATCTGGCCTCAAGTCCGCTCTCTCTCTCCCTGTTTACCAGG 156

Qy 118 AAGCAAAACCTTGACATCACTTCCATCTAGCAATGTTGGAAGAGTTAGTGCATGCGAG 177
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
157 AAACAAACCTTGACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 216

Qy 178 GTGTGGC 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 217 GTGTGGC 223

RESULT 5

US-11-247-587-127
; Sequence 127, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 127
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-127

Query Match 17.5%; Score 119.8; DB 8; Length 700;
Best Local Similarity 81.3%; Pred. No. 1.7e-30;
Matches 152; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

QY 1 ATGGCTTCTCTGTCATTTCTTCA---GCAGCTGTTGGCCACGCGCAGCAAGTTTACACAA 57
Db 2 ATGGCTTCTCTGTCATTTCTTCA---GCAGCTGTTGGCCACGCGCAGCAAGTTTACACAA 61

QY 58 GCTAGCATGGTGGCAGCTTTCACTGCTCAAAATCTTCAGCCACTTTTCCCTGTTACAAAG 117
Db 62 GCCAGTATGTTGCACCTTTCACTGGCTCAAGTCCGCAACCTCTTCCCTGTTTCCAGA 121

QY 118 AGCAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAG 177
Db 122 AAACAAACCTTGACATCACTTCCATGCTAGCAACGCGGAGAGTTCAATGTCATGCAG 181

QY 178 GTGTGGC 184
Db 182 GTGTGGC 188

RESULT 6

US-11-247-587-129
; Sequence 129, Application US/11247587
; Publication No. US20060130176A1
; GENERAL INFORMATION:
; APPLICANT: The Rockefeller University
; APPLICANT: Toboda, Jose Reyes
; APPLICANT: Zhang, Xiuren
; APPLICANT: Soyano, Takashi
; APPLICANT: Chua, Nam-Hai
; APPLICANT: Niu, Qi-Wen
; APPLICANT: Lin, Shih-Shun
; TITLE OF INVENTION: MicroRNAs
; FILE REFERENCE: 2312-124
; CURRENT APPLICATION NUMBER: US/11/247,587
; CURRENT FILING DATE: 2005-10-12
; PRIOR APPLICATION NUMBER: US 60/671,089
; PRIOR FILING DATE: 2005-04-14
; PRIOR APPLICATION NUMBER: PCT/US2004/033379
; PRIOR FILING DATE: 2004-10-12
; NUMBER OF SEQ ID NOS: 174

;; SOFTWARE: FastSeq for Windows Version 4.0/Patent-In 3.3
; SEQ ID NO 129
; LENGTH: 727
; TYPE: DNA
; ORGANISM: Nicotiana benthamiana
US-11-247-587-129

Query Match 16.1%; Score 109.8; DB 8; Length 727;
Best Local Similarity 78.4%; Pred. No. 4.5e-27;
Matches 145; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 3 GCCTTCTCTGTCATTTCTTCA---GCAGCTGTTGGCCACGCGCAGCAAGTTTACACAAGC 59
Db 7 GCCTTCTCTGTCATTTCTTCA---GCAGCTGTTGGCCACGCGCAGCAAGTTTACACAAGC 66

QY 60 TAGCATGGTGGCAGCTTTCACTGCTCAAAATCTTCAGCCACTTTTCCCTGTTTACAAGAA 119
Db 67 CAGCATGGTGGCAGCTTTCACTGCTCAAAATCTTCAGCCACTTTTCCCTGTTTACAAGAA 126

QY 120 GCAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGT 179
Db 127 ACAAAACCTTGACATTTCTTCACTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGT 186

QY 180 GTGGC 184
Db 187 GTGGC 191

RESULT 7

US-10-953-349-10521
; Sequence 10521, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10521
; LENGTH: 718
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-10521

Query Match 12.5%; Score 85.2; DB 6; Length 718;
Best Local Similarity 70.7%; Pred. No. 1.2e-18;
Matches 130; Conservative 0; Mismatches 48; Indels 6; Gaps 1;

QY 1 ATGGCTTCTCTGTCATTTCTTCAAGCAGCTGTTGGCCACGCGCAGCAAGTTTACACAAGCT 60
Db 32 ATGGCTTCTCTGTCATTTCTTCAAGCAGCTGTTGGCCACGCGCAGCAAGTTTACACAAGCT 85

QY 61 AGCATGGTGGCAGCTTTCACTGCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAGAAAG 120
Db 86 ACCATGGTGGCAGCTTTCACTGCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAGAAAG 145

QY 121 CAAAACCTTGACATCACTTTCCATTTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
Db 146 GCAACACAGCATTTACTTCCATCAACAGCAATGGGGAGAGTTAGCTGCATGCAGGTG 205

QY 181 TGGC 184
Db 206 TGGC 209

RESULT 8

US-11-218-305-17631
; Sequence 17631, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:

```
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17631
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17631

Query Match      12.2%; Score 83.6; DB 7; Length 583;
Best Local Similarity 70.1%; Pred. No. 3.9e-18;
Matches 129; Conservative 0; Mismatches 49; Indels 6; Gaps 1;

QY 1 ATGGCTTCCTCTGTCATTTCTTTCAGCAGCTGTTCACACGAGCAATGTTACACAAGCT 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 ATGGCTTCCTCTATGCTCTCTCCCGCGCTGTGGCTAC-----CTCACCGGCTCAAGCC 86
QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAATCTTCAGCCACTTTCCTCGTTTACAAAGAG 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
87 ACCATGGTTGCTCCATTACCGCGCTTGAAGTCAATCCGCTGCAATCCCAAGTACCCGCAAA 146
QY 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
147 ACACAACACTGATATTACTTCCATTGCAAGCAACGAGGAGAGTTAGCTGCATGAAGGTG 206
QY 181 TGGC 184
Db |||||
207 TGGC 210

RESULT 9
US-11-218-305-17633
; Sequence 17633, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; PRIOR FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17633
; LENGTH: 658
; TYPE: DNA
; ORGANISM: Zea mays
US-11-218-305-17633

Query Match      12.2%; Score 83.6; DB 7; Length 658;
Best Local Similarity 70.1%; Pred. No. 4.1e-18;
Matches 129; Conservative 0; Mismatches 49; Indels 6; Gaps 1;

QY 1 ATGGCTTCCTCTGTCATTTCTTTCAGCAGCTGTTCACACGAGCAATGTTACACAAGCT 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
45 ATGGCTTCCTCTATGCTTTCCTCCGCTGCTGTGGTTAC-----CTCCCCGGCTCAAGCC 98
QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAATCTTCAGCCACTTTCCTCGTTTACAAAGAG 120
```

```
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 ACCATGGTCGCTCCATTCACCGGCTTGAAGTCATCCGCTGCATTCCTCAGTCACCCGCAAA 158
QY 121 CAAAACCTTGACATCACTTTCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
159 GCAAACAACGACATTTACTTCCATTTGCAAGCAACGAGGAGAGTTAGCTGCATGAAGGTG 218
QY 181 TGGC 184
Db |||||
219 TGGC 222

RESULT 10
US-10-953-349-8894
; Sequence 8894, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8894
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-953-349-8894

Query Match      12.0%; Score 82; DB 6; Length 743;
Best Local Similarity 69.6%; Pred. No. 1.6e-17;
Matches 128; Conservative 0; Mismatches 50; Indels 6; Gaps 1;

QY 1 ATGGCTTCCTCTGTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
53 ATGGCTTCCTCTATGCTCTCTCCCGCGCTGTGGTTACATCC-----CCGGCTCAGGCC 106
QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAATCTTCAGCCACTTTCCTCGTTTACAAAGAG 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
107 ACCATGGTCGCTCCATTCACCGGCTTGAAGTCATCCGCTGCAATCCCGGTCAACCCGCAAG 166
QY 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 ACCAACAAGGACATCACTTCCATCGCAAGCAACGCGGGAAGAGTTAGCTGCATGAAGGTG 226
QY 181 TGGC 184
Db |||||
227 TGGC 230

RESULT 11
US-11-218-305-17634
; Sequence 17634, Application US/11218305
; Publication No. US20060141495A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO TECHNOLOGY, LLC
; APPLICANT: McLaird, Paul L.
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53660)B
; CURRENT APPLICATION NUMBER: US/11/218,305
; CURRENT FILING DATE: 2005-09-01
; PRIOR APPLICATION NUMBER: US 60/606,880
; PRIOR FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 25043
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17634
; LENGTH: 640
; TYPE: DNA
```

; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n is a, c, g, or t
US-11-218-305-17634

Query Match 11.8%; Score 80.4; DB 7; Length 640;
Best Local Similarity 69.0%; Pred. No. 5.1e-17;
Matches 127; Conservative 0; Mismatches 51; Indels 6; Gaps 1;
QY 1 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 60
Db 33 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 86
QY 61 AGCATGGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
Db 87 ACCATGGTGCCTCCATTACCGGCTTGAAGTCTTCCTGTCATTCCCAAGTCAACGCAAG 146
QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
Db 147 GCCAACCAACGACATTACCTCCATCGTTAGCAACGAGGAAGAGTTAGCTGCATGCAGGTG 206
QY 181 TGGC 184
Db 207 TGGC 210

RESULT 12
US-11-216-545-4174
; Sequence 4174, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; PRIOR FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4174
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: n is a, c, g, or t
US-11-216-545-4174

Query Match 11.5%; Score 78.4; DB 8; Length 1023;
Best Local Similarity 71.7%; Pred. No. 3.2e-16;
Matches 132; Conservative 0; Mismatches 46; Indels 6; Gaps 2;
QY 1 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 60
Db 228 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 284
QY 61 AGCATGGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
Db 285 GGCATGGTGTCTCCATTCACTGCGCTCAAGTCAATGCTGGCTTCCCC---ACGAGGAAG 341
QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
Db 342 ACCAACAAATGACATTACCTCCATTTGCTAGCAACGCGTGGAGAGTGAATGCAATGCATGCAGGTG 401
QY 181 TGGC 184

Db 402 TGGC 405

RESULT 13

US-11-216-545-4173
; Sequence 4173, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; PRIOR FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4173
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Glycine max
US-11-216-545-4173

Query Match 11.2%; Score 76.8; DB 8; Length 841;
Best Local Similarity 71.2%; Pred. No. 1e-15;
Matches 131; Conservative 0; Mismatches 47; Indels 6; Gaps 2;
QY 1 ATGGCTTCTCTGTCATTCTTCAGCAGCTGTTGCCACACGCGAGCAATGTTACACAAGCT 60
Db 37 ATGGCTTCTCTCAATGATCTCTCCAGCTGTTACCAC---CGTCAACCGTGGCGTGCC 93
QY 61 AGCATGGTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTGTTACAAAGAAG 120
Db 94 GGCATGGTGTCTCCATTACCGGCTCAAAATCCATGGCTGGCTTCCCC---ACGAGGAAG 150
QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
Db 151 ACCAACAAATGACATTACCTCCATTTGCTAGCAACGCGTGGAGAGTGAATGCAATGCATGCAGGTG 210
QY 181 TGGC 184
Db 211 TGGC 214

RESULT 14

US-11-216-545-8381
; Sequence 8381, Application US/11216545
; Publication No. US20060135758A1
; GENERAL INFORMATION:
; APPLICANT: MONSANTO Technology, LLC
; APPLICANT: McLaird, Paul L
; APPLICANT: Tao, Nengbing
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Single Nucleotide Polymorphic Markers and Methods of Genotyping
; FILE REFERENCE: 38-21 (53659)B
; CURRENT APPLICATION NUMBER: US/11/216,545
; PRIOR FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: US 60/606,062
; PRIOR FILING DATE: 2004-08-31
; NUMBER OF SEQ ID NOS: 8783
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8381
; LENGTH: 1237
; TYPE: DNA
; ORGANISM: Glycine max
US-11-216-545-8381

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 11, 2006, 19:18:20 ; Search time 1466 Seconds
(without alignments)
5733.112 Million cell updates/sec
Title: US-10-718-311-7
Perfect score: 684
Sequence: 1 atggcttctctgtcatttc.....cggcgtcaccgtgtactaa 684
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 18892170 seqs, 6143817638 residues
Total number of hits satisfying chosen parameters: 37784340
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				Published Applications NA Main:*			
				1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*			
				2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*			
				3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*			
				4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*			
				5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*			
				6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*			
				7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*			
				8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*			
				9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*			
				10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*			
				11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*			
				12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*			
				13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*			
				14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*			
				15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*			
				16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*			
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.							
SUMMARIES							
Result No.	Score	Query Match	Length	DB ID	Description		
1	684	100.0	684	3	US-09-855-341-7	Sequence 7, Appli	
2	684	100.0	684	3	US-09-896-866B-16	Sequence 16, Appl	
3	684	100.0	684	7	US-10-359-369-41	Sequence 41, Appl	
4	684	100.0	684	8	US-10-699-050-16	Sequence 16, Appl	
5	684	100.0	684	8	US-10-718-311-7	Sequence 7, Appli	
6	684	100.0	684	9	US-10-462-162-33	Sequence 33, Appl	
7	501	73.2	3452	9	US-10-462-162-47	Sequence 47, Appl	
8	499.4	73.0	1971	9	US-10-462-162-54	Sequence 54, Appl	
9	498	72.8	498	9	US-10-462-162-29	Sequence 29, Appl	
10	495	72.4	495	3	US-09-855-341-3	Sequence 3, Appli	
11	495	72.4	495	3	US-09-896-866B-12	Sequence 12, Appl	
12	495	72.4	495	7	US-10-359-369-37	Sequence 37, Appl	
13	495	72.4	495	8	US-10-699-050-12	Sequence 12, Appl	
14	495	72.4	495	8	US-10-718-311-3	Sequence 3, Appli	
15	472.8	69.1	1207	10	US-10-450-763-26139	Sequence 26139, A	
16	144	21.1	718	10	US-10-487-901-5765	Sequence 5765, Ap	
17	144	21.1	736	10	US-10-487-901-1704	Sequence 1704, Ap	

18	144	21.1	736	10	US-10-487-901-5043	Sequence 5043, Ap
19	142.4	20.8	297	10	US-10-487-901-7209	Sequence 7209, Ap
20	142.4	20.8	377	10	US-10-487-901-3504	Sequence 3504, Ap
21	142.4	20.8	489	10	US-10-487-901-3503	Sequence 3503, Ap
22	142.4	20.8	504	10	US-10-487-901-7207	Sequence 7207, Ap
23	140.8	20.6	204	3	US-09-839-477-3	Sequence 3, Appli
24	140.8	20.6	204	8	US-10-758-064-3	Sequence 3, Appli
25	137.6	20.1	668	10	US-10-487-901-7208	Sequence 7208, Ap
26	129.2	18.9	174	3	US-09-854-286-15	Sequence 15, Appl
27	128.8	18.8	169	6	US-10-165-420-3	Sequence 3, Appli
28	124.6	18.2	683	10	US-10-487-901-1713	Sequence 1713, Ap
29	124.6	18.2	737	10	US-10-487-901-5045	Sequence 5045, Ap
30	124.6	18.2	737	10	US-10-487-901-5766	Sequence 5766, Ap
31	121.4	17.7	738	10	US-10-487-901-1716	Sequence 1716, Ap
32	121.4	17.7	753	10	US-10-487-901-5764	Sequence 5764, Ap
33	121.4	17.7	754	10	US-10-487-901-5028	Sequence 5028, Ap
34	119.8	17.5	714	10	US-10-487-901-1703	Sequence 1703, Ap
35	119.8	17.5	736	10	US-10-487-901-5044	Sequence 5044, Ap
36	115	16.8	704	9	US-10-425-115-85509	Sequence 85509, A
37	113	16.5	151	7	US-10-321-434-4	Sequence 4, Appli
38	111.8	16.3	654	10	US-10-487-901-5042	Sequence 5042, Ap
39	111	16.2	180	10	US-10-487-901-88	Sequence 88, Appl
40	97.2	14.2	149	3	US-09-839-477-5	Sequence 5, Appli
41	97.2	14.2	149	8	US-10-758-064-5	Sequence 5, Appli
42	85.2	12.5	363	3	US-09-770-791-468	Sequence 468, App
43	85.2	12.5	711	3	US-09-910-664-32	Sequence 32, Appl
44	85.2	12.5	711	8	US-10-333-184-31	Sequence 31, Appl
45	85.2	12.5	739	8	US-10-333-184-182	Sequence 182, App

ALIGNMENTS

RESULT 1
US-09-855-341-7
; Sequence 7, Application US/09855341
; Patent No. US20020002715A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUST
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-09-855-341-7

Query Match		100.0%	Score 684;	DB 3;	Length 684;
Best Local Similarity		100.0%	Pred. No. 1.9e-222;		
Matches 684;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT	60		
DB	1	ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTCACACGCGAGCAATGTTACACAAGCT	60		
QY	61	AGCATGGTTGACCTTTCTCAATCTTTCAGCCACTTTTCCCTGTTACAAAGAAG	120		
DB	61	AGCATGGTTGACCTTTCTCAATCTTTCAGCCACTTTTCCCTGTTACAAAGAAG	120		
QY	121	CAAAACCTTGACATCACTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG	180		
DB	121	CAAAACCTTGACATCACTTCCATTCCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG	180		
QY	181	TGGCATATGTCACACCCCGGTTACGCAACTTCGCTGCGCTATTGTTAAAGAGATC	240		

```
Db 181 TGGCATATGTACACCCCGGTTAAGCNACTGCGTGGCTAGCGCTATTGTAAGAGATC 240
Qy 241 CCTGCCCTGGATCCGGAACCTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300
Db 241 CCTGCCCTGGATCCGGAACCTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300
Qy 301 TTTGAACAGCAGGGAAGAAACGGTAAGCGTGACGATATCCGGAAGGGTTTGTCCGAGCAG 360
Db 301 TTTGAACAGCAGGGAAGAAACGGTAAGCGTGACGATATCCGGAAGGGTTTGTCCGAGCAG 360
Qy 361 AATGAATCCCGGAAGAACTGCCGTGCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAA 420
Db 361 AATGAATCCCGGAAGAACTGCCGTGCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAA 420
Qy 421 ATTTTGTATGTGCCGATGGTGAAACGGTGGCTTGGCCGTTACCGTCTGTTCTCTGTGCA 480
Db 421 ATTTTGTATGTGCCGATGGTGAAACGGTGGCTTGGCCGTTACCGTCTGTTCTCTGTGCA 480
Qy 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540
Qy 541 CTGTTTACATCATGACATTAACCGGGGACTTTTATTGAGATAGGCCGCTGATGCCGGGCTG 600
Db 541 CTGTTTACATCATGACATTAACCGGGGACTTTTATTGAGATAGGCCGCTGATGCCGGGCTG 600
Qy 601 TGGGGGCGACGTTCCCGCTGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGACGTTCCCGCTGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy 661 TTACCGGCGTACCGTTTGACTAA 684
Db 661 TTACCGGCGTACCGTTTGACTAA 684

RESULT 2
US-09-896-866B-16
; Sequence 16, Application US/09896866B
; Patent No. US20020151002A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose-Malate Sinapoyltransferase Form Malate Conjugates
; TITLE OF INVENTION: Benzoic Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896,866B
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protein
US-09-896-866B-16
```

Query Match 100.0%; Score 684; DB 3; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.9e-222;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGGCTTCTCTGTCTATTCTTTACAGAGCTGTTCACAGCAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCTCTGTCTATTCTTTACAGAGCTGTTCACAGCAGCAATGTTACACAAGCT 60
Qy 61 AGCATGGTGGACCTTTCACCTGCTCAAACTTCAGCCACTTTCCCTGTTACAAAGAG 120
Db 61 AGCATGGTGGACCTTTCACCTGCTCAAACTTCAGCCACTTTCCCTGTTACAAAGAG 120
```

```
Qy 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGCTGGAAGAGTTAGCTGCATGCAAGTG 180
Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGCTGGAAGAGTTAGCTGCATGCAAGTG 180
Qy 181 TGGCATATGTACACCCCGGTTAAGCNACTGCGTGGCTAGCGCTATTGTAAGAGATC 240
Db 181 TGGCATATGTACACCCCGGTTAAGCNACTGCGTGGCTAGCGCTATTGTAAGAGATC 240
Qy 241 CTGCCCTGGATCCGGAACCTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300
Db 241 CTGCCCTGGATCCGGAACCTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAAAGT 300
Qy 301 TTTGAACAGCAGGGAAGAAACGGTAAGCGTGACGATATCCGGAAGGGTTTGTCCGAGCAG 360
Db 301 TTTGAACAGCAGGGAAGAAACGGTAAGCGTGACGATATCCGGAAGGGTTTGTCCGAGCAG 360
Qy 361 AATGAATCCCGGAAGAACTGCCGTGCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAA 420
Db 361 AATGAATCCCGGAAGAACTGCCGTGCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAA 420
Qy 421 ATTTTGTATGTGCCGATGGTGAAACGGTGGCTTGGCCGTTACCGTCTGTTCTCTGTGCA 480
Db 421 ATTTTGTATGTGCCGATGGTGAAACGGTGGCTTGGCCGTTACCGTCTGTTCTCTGTGCA 480
Qy 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACGTTAAGCGGCGCGAGCTGGCGTTACAAAAATTTGGGTAAACCGCGTTAGGACGCTAT 540
Qy 541 CTGTTTACATCATGACATTAACCGGGGACTTTTATTGAGATAGGCCGCTGATGCCGGGCTG 600
Db 541 CTGTTTACATCATGACATTAACCGGGGACTTTTATTGAGATAGGCCGCTGATGCCGGGCTG 600
Qy 601 TGGGGGCGACGTTCCCGCTGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGACGTTCCCGCTGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy 661 TTACCGGCGTACCGTTTGACTAA 684
Db 661 TTACCGGCGTACCGTTTGACTAA 684
```

RESULT 3
US-10-359-369-41
; Sequence 41, Application US/10359369
; Publication No. US20030215927A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Viitanen, Paul
; APPLICANT: Meyer, Knut
; APPLICANT: Van Dyk, Drew

; TITLE OF INVENTION: UDP-Glucosyltransferases
; FILE REFERENCE: CL1821 US NA
; CURRENT APPLICATION NUMBER: US/10/359,369
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 60/355,511
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 41
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(684)
US-10-359-369-41

Query Match 100.0%; Score 684; DB 7; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.9e-222;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCTTCTCTGTCTATTCTTTCAGCAGCTGTTGCCACACGCAATGTTACACAAGCT 60


```
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic CPL
US-10-718-311-7
```

```
Query Match      100.0%; Score 684; DB 8; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.9e-222;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60
DB 1 ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60

QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAAG 120
DB 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAAG 120

QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180

QY 181 TGGCATATGTCAACCCCGCGTTAAACGCAACTGCGTGGCGTCTATTGTTAAAGAGATC 240
DB 181 TGGCATATGTCAACCCCGCGTTAAACGCAACTGCGTGGCGTCTATTGTTAAAGAGATC 240

QY 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300
DB 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300

QY 301 TTTGACAGCAGGGAACGTTAGCGTGACGATGATCCGGAAGGGTTTGTCCGAGCAG 360
DB 301 TTTGACAGCAGGGAACGTTAGCGTGACGATGATCCGGAAGGGTTTGTCCGAGCAG 360

QY 361 AATGAAATCCCGGAAGAACTGCGGTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAA 420
DB 361 AATGAAATCCCGGAAGAACTGCGGTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAA 420

QY 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTTGGCGGTGCGTACCGTCTGCTGTGTCA 480
DB 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTTGGCGGTGCGTACCGTCTGCTGTGTCA 480

QY 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTAT 540
DB 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTAT 540

QY 541 CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600
DB 541 CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600

QY 601 TGGGGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
DB 601 TGGGGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660

QY 661 TTACCGGCGTCAACCGTTGTACTAA 684
DB 661 TTACCGGCGTCAACCGTTGTACTAA 684
```

```
RESULT 6
US-10-462-162-33
; Sequence 33, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
; CURRENT APPLICATION NUMBER: US/10/462,162
```

```
; CURRENT FILING DATE: 2003-06-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 33
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-462-162-33
```

```
Query Match      100.0%; Score 684; DB 9; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.9e-222;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60
DB 1 ATGGCTTCTCTGTCTATTTCTTTCAGCAGCTGTTCACACGACGCAATGTTACACAAGCT 60

QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAAG 120
DB 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAAG 120

QY 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
DB 121 CAAAACCTTGACATCACTTCCATTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180

QY 181 TGGCATATGTCAACCCCGCGTTAAACGCAACTGCGTGGCGTCTATTGTTAAAGAGATC 240
DB 181 TGGCATATGTCAACCCCGCGTTAAACGCAACTGCGTGGCGTCTATTGTTAAAGAGATC 240

QY 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300
DB 241 CCTGCCCTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGT 300

QY 301 TTTGACAGCAGGGAACGTTAGCGTGACGATGATCCGGAAGGGTTTGTCCGAGCAG 360
DB 301 TTTGACAGCAGGGAACGTTAGCGTGACGATGATCCGGAAGGGTTTGTCCGAGCAG 360

QY 361 AATGAAATCCCGGAAGAACTGCGGTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAA 420
DB 361 AATGAAATCCCGGAAGAACTGCGGTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAA 420

QY 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTTGGCGGTGCGTACCGTCTGCTGTGTCA 480
DB 421 ATTTTGTATGTGCGCATGGTGAACCGTGGCTTGGCGGTGCGTACCGTCTGCTGTGTCA 480

QY 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTAT 540
DB 481 ACGTTAAGCGGCGCGAGCTGCGTTACAAAATTTGGGTAAACCGCTTAGGACGCTAT 540

QY 541 CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600
DB 541 CTGTTTCACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600

QY 601 TGGGGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
DB 601 TGGGGGCGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660

QY 661 TTACCGGCGTCAACCGTTGTACTAA 684
DB 661 TTACCGGCGTCAACCGTTGTACTAA 684
```

```
RESULT 7
US-10-462-162-47
; Sequence 47, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; APPLICANT: Flint, Dennis
; TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
; FILE REFERENCE: CL 2155 US NA
```



```

; ORGANISM: artificial sequence
;
; FEATURE:
;
; OTHER INFORMATION: Nucleic acid sequence of a nucleic acid fragment inserted into
;
; OTHER INFORMATION: expression vector pET29a encoding CPL (from Escherichia coli) an
;
; OTHER INFORMATION: pHBA 1-hydroxylase (from Candida parapsilosis).
US-10-462-162-54

```

Query Match . 73.0%; Score 499.4; DB 9; Length 1971;
Best Local Similarity 99.8%; Pred. No. 4.6e-159;

		Matches	500;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	184	CATATGTCACACCCCGGTTAAACGCAACTGCGTGCCTGCGCTATTGTAAAGAGATCCCT	243								
Db	1	CATATGTCACACCCCGGTTAAACGCAACTGCGTGCCTGCGCTATTTTAAAGAGATCCCT	60								
Qy	244	GCCTTGATCCGGAACCTGCTCGACTGCGTGTGCTGGAGGATTCCATGACAAAAACGTTTT	303								
Db	61	GCCTTGATCCGGAACCTGCTCGACTGCGTGTGCTGGAGGATTCCATGACAAAAACGTTTT	120								
Qy	304	GAACAGCAGGGAAGAAACGGTAAGCGTCACCATGATCCGCGAAGGGTTTGTCCAGCAGAAAT	363								
Db	121	GAACAGCAGGGAAGAAACGGTAAGCGTCACCATGATCCGCGAAGGGTTTGTCCAGCAGAAAT	180								
Qy	364	GAATCCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAATTT	423								
Db	181	GAATCCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAATTT	240								
Qy	424	TTGTTATGTGCCGATGGTGAACCGTGGCTTTGCCGGTGGTACCGTGGTTCTGTGTCAACG	483								
Db	241	TTGTTATGTGCCGATGGTGAACCGTGGCTTTGCCGGTGGTACCGTGGTTCTGTGTCAACG	300								
Qy	484	TTAAGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAAACCGCCGTTAGGACGCTATCTG	543								
Db	301	TTAAGCGGGCGGAGCTGGCGTTACAAAAATTGGGTAAACCGCCGTTAGGACGCTATCTG	360								
Qy	544	TTCAACATCATCGACATTAAACCGGAGCTTTATTGAGATAGCCGCTGATGCCGGGCTGGG	603								
Db	361	TTCAACATCATCGACATTAAACCGGAGCTTTATTGAGATAGCCGCTGATGCCGGGCTGGG	420								
Qy	604	GGCGAGCGTTCCCGCCTCGCAATTAAACGGGTAAACCGCTGTTGTCTAAACAGAACTGTTTTTA	663								
Db	421	GGCGAGCGTTCCCGCCTCGCAATTAAACGGGTAAACCGCTGTTGTCTAAACAGAACTGTTTTTA	480								
Qy	664	CCGGGCTCACCGTTGACTAA	684								
Db	481	CCGGGCTCACCGTTGACTAA	501								

RESULT 9
US-10-462-162-29
; Sequence 29, Application US/10462162
; Publication No. US20040261147A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.

```

/ APPLICANT: Meyer, Anne
/ APPLICANT: Viitanen, Paul
/ APPLICANT: Flint, Dennis
/ TITLE OF INVENTION: High Level Production of Arbutin in Green Plants and Microbes
/ FILE REFERENCE: CL 2155 US NA
/ CURRENT APPLICATION NUMBER: US/10/462,162
/ CURRENT FILING DATE: 2003-06-16
/ NUMBER OF SEQ ID NOS: 54
/ SOFTWARE: Microsoft Office 97

```

```
Query Match      72.8%; Score 498; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.9e+159;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTCTGTAACCGTTCCTTCTGTGTCAACGTTA 300
QY 487 AGCGGCCGAGGCTGGCGTTACAAAAATTGGGTAAAAACCGGTTAGGACGCTATCTGTTC 546
Db 301 AGCGGCCGAGGCTGGCGTTACAAAAATTGGGTAAAAACCGGTTAGGACGCTATCTGTTC 360
QY 547 ACATCATCGACATTAACCGGAGCTTTATTAGATAGGCGGTGATCGCGGCTGTGGGG 606
Db 361 ACATCATCGACATTAACCGGAGCTTTATTAGATAGGCGGTGATCGCGGCTGTGGGG 420
QY 607 CGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 666
Db 421 CGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 480
QY 667 GCGTCAACCGTTGTAC 681
Db 481 GCGTCAACCGTTGTAC 495

RESULT 12

US-10-359-369-37
; Sequence 37, Application US/10359369
; Publication No. US20030215927A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Viitanen, Paul
; APPLICANT: Meyer, Knut
; APPLICANT: Van Dyk, Drew
; TITLE OF INVENTION: UDP-Glucosyltransferases
; FILE REFERENCE: CL1821 US NA
; CURRENT APPLICATION NUMBER: US/10/359,369
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: 60/355,511
; PRIOR FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(495)
US-10-359-369-37

Query Match 72.4%; Score 495; DB 7; Length 495;
Best Local Similarity 100.0%; Pred. No. 7.3e-158;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 ATGTACACCCCGCGTTAACGCACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 246
Db 1 ATGTACACCCCGCGTTAACGCACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 60
QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCCATGACAAAACGTTTGA 306
Db 61 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCCATGACAAAACGTTTGA 120
QY 307 CAGCAGGAAAAACCGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 366
Db 121 CAGCAGGAAAAACCGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 180
QY 367 ATCCCGGAAGAACTGCGCTGCTGCGGAAAGAGTCTCGTTACTGTTAGTAAATTTG 426
Db 181 ATCCCGGAAGAACTGCGCTGCTGCGGAAAGAGTCTCGTTACTGTTAGTAAATTTG 240
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTCTGCTACCGTTCCTGTGTCAACGTTA 486
Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTCTGCTACCGTTCCTGTGTCAACGTTA 300
QY 487 AGCGGCCGAGGCTGGCGTTACAAAAATTGGGTAAAAACCGGTTAGGACGCTATCTGTTC 546
Db 301 AGCGGCCGAGGCTGGCGTTACAAAAATTGGGTAAAAACCGGTTAGGACGCTATCTGTTC 360

QY 547 ACATCATCGACATTAACCGGAGCTTTATTAGATAGGCGGTGATCGCGGCTGTGGGG 606
Db 361 ACATCATCGACATTAACCGGAGCTTTATTAGATAGGCGGTGATCGCGGCTGTGGGG 420
QY 607 CGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 666
Db 421 CGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 480
QY 667 GCGTCAACCGTTGTAC 681
Db 481 GCGTCAACCGTTGTAC 495

RESULT 13

US-10-699-050-12
; Sequence 12, Application US/10699050
; Publication No. US20040142437A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Dennis
; APPLICANT: Meyer, Knut
; APPLICANT: Viitanen, Paul
; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/10/699,050
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US/09/896,866B
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
US-10-699-050-12

Query Match 72.4%; Score 495; DB 8; Length 495;
Best Local Similarity 100.0%; Pred. No. 7.3e-158;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 ATGTACACCCCGCGTTAACGCACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 246
Db 1 ATGTACACCCCGCGTTAACGCACTGCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 60
QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCCATGACAAAACGTTTGA 306
Db 61 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATTCCATGACAAAACGTTTGA 120
QY 307 CAGCAGGAAAAACCGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 366
Db 121 CAGCAGGAAAAACCGTAAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 180
QY 367 ATCCCGGAAGAACTGCGCTGCTGCGGAAAGAGTCTCGTTACTGTTAGTAAATTTG 426
Db 181 ATCCCGGAAGAACTGCGCTGCTGCGGAAAGAGTCTCGTTACTGTTAGTAAATTTG 240
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTCTGCTACCGTTCCTGTGTCAACGTTA 486
Db 241 TTATGTGCCGATGGTGAACCGTGGCTTGGCGGTCTGCTACCGTTCCTGTGTCAACGTTA 300
QY 487 AGCGGCCGAGGCTGGCGTTACAAAAATTGGGTAAAAACCGGTTAGGACGCTATCTGTTC 546
Db 301 AGCGGCCGAGGCTGGCGTTACAAAAATTGGGTAAAAACCGGTTAGGACGCTATCTGTTC 360
QY 547 ACATCATCGACATTAACCGGAGCTTTATTAGATAGGCGGTGATCGCGGCTGTGGGG 606
Db 361 ACATCATCGACATTAACCGGAGCTTTATTAGATAGGCGGTGATCGCGGCTGTGGGG 420
QY 607 CGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 666
Db 421 CGACGTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTGCTAACAGAACTGTTTTACCG 480

Qy 667 GCGTACCGTGTGAC 681
Db 481 GCGTACCGTGTGAC 495

RESULT 14

US-10-718-311-3
; Sequence 3, Application US/10718311
; Publication No. US20040143867A1
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUIT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/10/718,311
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/855,341
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: E. coli
US-10-718-311-3

Query Match 72.4%; Score 495; DB 8; Length 495;
Best Local Similarity 100.0%; Pred. No. 7.3e-158;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 187 ATGTACACACCCCGGTTAACGCAACTGCGTGGCTGCGCTATTGTGAAAGAGATCCCTGCC 246
Db 1 ATGTACACACCCCGGTTAACGCAACTGCGTGGCTGCGCTATTGTGAAAGAGATCCCTGCC 60

Qy 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATCCATGACAAAAACGTTTGA 306
Db 61 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATCCATGACAAAAACGTTTGA 120

Qy 307 CAGCAGGGAACAAACGGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGAAATGA 366
Db 121 CAGCAGGGAACAAACGGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGAAATGA 180

Qy 367 ATCCCGGAAGAACTCCCGTGTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTTG 426
Db 181 ATCCCGGAAGAACTCCCGTGTGCTGCGAAGAGTCTCGTTACTGTTGAAATTTTG 240

Qy 427 TTATGTCCGATGGTGAACCGTGGCTTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
Db 241 TTATGTCCGATGGTGAACCGTGGCTTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

Qy 487 AGCGGCGCGAGCTGGCGTTACAAAAATTGGTAAACCGCGTTAGGACGCTATCTGTTTC 546
Db 301 AGCGGCGCGAGCTGGCGTTACAAAAATTGGTAAACCGCGTTAGGACGCTATCTGTTTC 360

Qy 547 ACATCATCGACATTAACCCCGGACCTTTATTTAGATAGCGCGTGTGTCGGGGCTGTGGGG 606
Db 361 ACATCATCGACATTAACCCCGGACCTTTATTTAGATAGCGCGTGTGTCGGGGCTGTGGGG 420

Qy 607 CGAGCTTCCCGCTCGAGTTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTCACG 666
Db 421 CGAGCTTCCCGCTCGAGTTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTCACG 480

Qy 667 GCGTACCGTGTGAC 681
Db 481 GCGTACCGTGTGAC 495

RESULT 15

US-10-450-763-26139
; Sequence 26139, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 26139
; LENGTH: 1207
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (893)..(925)
; OTHER INFORMATION: 81% homologous to Escherichia coli 4-hydroxybenzoate
; OTHER INFORMATION: synthetase, accession number X57434, Smith-Waterman Score=49.
US-10-450-763-26139

Query Match 69.1%; Score 472.8; DB 10; Length 1207;
Best Local Similarity 99.2%; Pred. No. 4.5e-150;
Matches 496; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 187 ATGTACACACCCCGGTTAACGCAACTGCGTGGCTGCGCTATTGTGAAAGAGATCCCTGCC 246
Db 652 ATGTACACACCCCGGTTAACGCAACTGCGTGGCTGCGCTATTGTGAAAGAGATCCCTGCC 711

Qy 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATCCATGACAAAAACGTTTGA 306
Db 712 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGGATCCATGACAAAAACGTTTGA 771

Qy 307 CAGCAGGGAACAAACGGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGAAATGA 366
Db 772 CAGCAGGGAACAAACGGTAAAGCGTACGATGATCCCGAAGGTTTGTCCAGCAGAAATGA 831

Qy 367 ATCCCGGAAGAACTCCCGTGTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTTG 426
Db 832 ATCCCGGAAGAACTCCCGTGTGCTGCGAAGAGTCTCGTTACTGTTACGTTGAAATTTTG 891

Qy 427 T-TATGTCCGATGGTGAACCGTGGCTTGCC-GGTGCTACCGTCTGTTCTCTGTCAACGT 484
Db 892 TGTATGTCCGATGGCGAAACCGCGGCTTGCCGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951

Qy 485 TAACCGGCGCGAGCTGGCGTTACAAAAATTGGGTAAACCGCGTTAGGACGCTATCTGT 544
Db 952 TAACCGGCGCGAGCTGGCGTTACAAAAATTGGGTAAACCGCGTTAGGACGCTATCTGT 1011

Qy 545 TCACATCATCGACATTAACCCCGGACCTTTATTTAGATAGGCGCGTATGTCGGGGCTGTGGG 604
Db 1012 TCACATCATCGACATTAACCCCGGACCTTTATTTAGATAGGCGCGTATGTCGGGGCTGTGGG 1071

Qy 605 GCGACGCTTCCCGCTCGAGTTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTCAC 664
Db 1072 GCGACGCTTCCCGCTCGAGTTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTCAC 1131

Qy 665 CGCGCTCACCGTGTGACTAA 684
Db 1132 CGCGCTCACCGTGTGACTAA 1151

Search completed: July 11, 2006, 19:46:02
Job time : 1467 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 19:14:52 ; Search time 190 Seconds
(without alignments)
6735.992 Million cell updates/sec

Title: US-10-718-311-7
Perfect score: 684
Sequence: 1 acggttctctgtctatttc.....cggcgtaccgttgtactaa 684

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /EMC_Celerra_SIDS3/ptodata/2/ina/1 COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5 COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7 COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	684	100.0	684	3	US-09-896-866B-16
2	684	100.0	684	3	US-09-855-341-7
3	498	72.8	854	3	US-09-064-693A-24
4	498	72.8	6641	3	US-09-064-693A-25
5	495	72.4	495	3	US-09-896-866B-12
6	495	72.4	495	3	US-09-855-341-3
7	282	41.2	645	3	US-09-489-039A-1234
8	140.8	20.6	204	3	US-09-839-477-3
9	140.8	20.6	499	3	US-09-495-797-41
10	133.2	19.5	1442	2	US-08-152-483B-8
11	130.6	19.1	177	2	US-08-095-726-12
12	130.6	19.1	177	2	US-08-096-043-12
13	130.6	19.1	177	2	US-08-093-577-8
14	130.6	19.1	177	2	US-08-331-004A-6
15	130.6	19.1	177	2	US-08-096-623A-12
16	130.6	19.1	177	2	PCT-US95-13937A-6
17	126.6	18.5	537	3	US-09-543-681A-3122
18	97.2	14.2	149	3	US-09-839-477-5
19	77.2	11.3	281	2	US-08-469-202-15
20	77.2	11.3	281	2	US-08-469-202-17
21	77.2	11.3	281	2	US-08-469-202-19
22	77.2	11.3	281	2	US-08-484-434C-15
23	77.2	11.3	281	2	US-08-484-434C-18

24	77.2	11.3	281	2	US-08-484-434C-22	Sequence 22, Appl
25	77.2	11.3	281	3	US-09-384-361-15	Sequence 15, Appl
26	77.2	11.3	281	3	US-09-384-361-18	Sequence 18, Appl
27	77.2	11.3	281	3	US-09-384-361-22	Sequence 22, Appl
28	75.2	11.0	104	2	US-08-095-726-38	Sequence 38, Appl
29	75.2	11.0	104	2	US-08-096-043-35	Sequence 35, Appl
30	75.2	11.0	104	2	US-08-093-577-31	Sequence 31, Appl
31	75.2	11.0	104	2	US-08-096-623A-43	Sequence 43, Appl
32	74.4	10.9	264	3	US-09-441-340-9	Sequence 9, Appl
33	74.4	10.9	268	3	US-09-186-002-5	Sequence 5, Appl
34	74.4	10.9	279	2	US-08-391-339-9	Sequence 9, Appl
35	74.4	10.9	279	2	US-08-484-274A-9	Sequence 9, Appl
36	74.4	10.9	279	2	US-09-612-404-9	Sequence 9, Appl
37	74.4	10.9	355	2	US-08-090-523-5	Sequence 5, Appl
38	74.4	10.9	355	2	US-08-398-627-5	Sequence 5, Appl
39	74.4	10.9	355	2	US-08-406-858-5	Sequence 5, Appl
40	74.4	10.9	355	2	US-08-476-519-7	Sequence 7, Appl
41	74.4	10.9	355	3	US-08-120-703A-5	Sequence 5, Appl
42	74.4	10.9	355	3	US-08-399-023-5	Sequence 5, Appl
43	74.4	10.9	355	7	PCT-US91-04036-5	Sequence 5, Appl
44	74.4	10.9	355	7	PCT-US94-05275-5	Sequence 5, Appl
45	74.4	10.9	355	7	PCT-US95-09323-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-896-866B-16

; Sequence 16, Application US/09896866B

; Patent No. 6642036

; GENERAL INFORMATION:

; APPLICANT: Flint, Dennis

; APPLICANT: Meyer, Knut

; APPLICANT: Viitanen, Paul

; TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate

; FILE REFERENCE: BC1034 US NA

; CURRENT APPLICATION NUMBER: US/09/896,866B

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 60/216,615

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16

; TYPE: DNA

; LENGTH: 684

; ORGANISM: Artificial Sequence

; NAME/KEY: misc_feature

; OTHER INFORMATION: open reading frame of the chloroplast-targeted CPL fusion protei

US-09-896-866B-16

Query Match 100.0%; Score 684; DB 3; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.1e-230;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTCTGTGCTATTTCCTTTCAGCAGCTGTTCACACAGCAGCAATGTACCAAGCT 60

Db 1 ATGGCTTCTCTGTGCTATTTCCTTTCAGCAGCTGTTCACACAGCAGCAATGTACCAAGCT 60

QY 61 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTCTGTACCAAGAAG 120

Db 61 AGCATGGTTGCACCTTTCACCTGCTCAATCTTCAGCCACTTTCCTCTGTACCAAGAAG 120

QY 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180

QY 181 TGGCATATGTCAACCCCGGTTTAACGAACTGCGTGCCTGCGCTATTGTTAAAGAGATC 240

Db 181 TGGCATATGTCAACCCCGGTTTAACGAACTGCGTGCCTGCGCTATTGTTAAAGAGATC 240

```
QY 241 CCTGCTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300
Db 241 CCTGCTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300
QY 301 TTTGAACAGCAGGAGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAG 360
Db 301 TTTGAACAGCAGGAGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAG 360
QY 361 AATGAATCCCGGAGAACTGCGCTGCTGCGAAGAGTCTCGTTACTGTTAGCTGAA 420
Db 361 AATGAATCCCGGAGAACTGCGCTGCTGCGAAGAGTCTCGTTACTGTTAGCTGAA 420
QY 421 ATTTTGTATGTGCGGATGTTGAACCGTGGCTTGCCTGCTGACCGTGGTCTCTGTGCA 480
Db 421 ATTTTGTATGTGCGGATGTTGAACCGTGGCTTGCCTGCTGACCGTGGTCTCTGTGCA 480
QY 481 AGCTTAAGCGGCGGAGCTGCGGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540
Db 481 AGCTTAAGCGGCGGAGCTGCGGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540
QY 541 CTGTTTACATCATCGACATTACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600
Db 541 CTGTTTACATCATCGACATTACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600
QY 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
QY 661 TTACCGGGCTACCGTTGTACTAA 684
Db 661 TTACCGGGCTACCGTTGTACTAA 684
```

RESULT 2

```
US-09-855-341-7
; Sequence 7, Application US/09855341
; Patent No. 6683231
; GENERAL INFORMATION:
; APPLICANT: MEYER, KNUZ
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; FILE REFERENCE: IN GREEN PLANTS
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 7
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic CPL
US-09-855-341-7
```

```
Query Match 100.0%; Score 684; DB 3; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.1e-230;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCTGTGCTATTTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCTGTGCTATTTTTCAGCAGCTGTTGCCACACGAGCAATGTTACACAAGCT 60
QY 61 AGCATGTTGACCTTTCAGCTGCTCAAACTTTCAGCCACTTTCCTGTTTACAAAGAG 120
Db 61 AGCATGTTGACCTTTCAGCTGCTCAAACTTTCAGCCACTTTCCTGTTTACAAAGAG 120
QY 121 CAAACCTTGACATCAGCTTTCATGCTAGCAATGTTGGAAGAGTTAGCTGATGAGGTG 180
Db 121 CAAACCTTGACATCAGCTTTCATGCTAGCAATGTTGGAAGAGTTAGCTGATGAGGTG 180
QY 181 TGGCATATGTACACCCCGGCTTAAACCAACTGCGTGGCTGCTATTTGTAAGAGATC 240
```

```
Db 181 TGGCATATGTACACCCCGGCTTAAACCAACTGCGTGGCTGCTATTTGTAAGAGATC 240
QY 241 CCTGCTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300
Db 241 CCTGCTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTTCATGACAAAACGT 300
QY 301 TTTGAACAGCAGGAGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAG 360
Db 301 TTTGAACAGCAGGAGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAG 360
QY 361 AATGAATCCCGGAGAACTGCGCTGCTGCGAAGAGTCTCGTTACTGTTAGCTGAA 420
Db 361 AATGAATCCCGGAGAACTGCGCTGCTGCGAAGAGTCTCGTTACTGTTAGCTGAA 420
QY 421 ATTTTGTATGTGCGGATGTTGAACCGTGGCTTGCCTGCTGACCGTGGTCTCTGTGCA 480
Db 421 ATTTTGTATGTGCGGATGTTGAACCGTGGCTTGCCTGCTGACCGTGGTCTCTGTGCA 480
QY 481 ACCTTAAGCGGCGGAGCTGCGGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACCTTAAGCGGCGGAGCTGCGGTTACAAAATTGGGTAAACCGCGTTAGGACGCTAT 540
QY 541 CTGTTTACATCATCGACATTAAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600
Db 541 CTGTTTACATCATCGACATTAAACCCGGGACTTTATTGAGATAGGCGGTGATGCCGGGCTG 600
QY 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGACGTTCCCGGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
QY 661 TTACCGGGCTACCGTTGTACTAA 684
Db 661 TTACCGGGCTACCGTTGTACTAA 684
```

RESULT 3

```
US-09-064-693A-24
; Sequence 24, Application US/09064693A
; Patent No. 6210937
; GENERAL INFORMATION:
; APPLICANT: Ward, Thomas E.
; TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
; ENGINEERED BACTERIA FOR PRODUCTION
; OF A SPECIFIC PLASTICS PRECURSOR
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Gary Goodson
; ADDRESS: INEL--Lockheed Martin Idaho
; ADDRESS: Technologies Co.
; STREET: P.O. Box 1625
; CITY: Idaho Falls
; STATE: Idaho
; COUNTRY: USA
; ZIP: 83415-3810
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: Toshiba Satellite Pro T2150CDS
; OPERATING SYSTEM: Windows95
; SOFTWARE: Word Perfect 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,693A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: W. Gary Goodson
; REGISTRATION NUMBER: 22,387
; REFERENCE/DOCKET NUMBER: LIT-PI-296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (208)526-9469
```

TELEFAX: (208)526-8339
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 854 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-064-693A-24

Query Match 72.8%; Score 498; DB 3; Length 854;
Best Local Similarity 100.0%; Pred. No. 6.2e-165;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 ATGTACACCCCGGTTAAGCAACTGCGTGCCTGCTATTTGTAAGAGATCCCTGCC 246
Db 321 ATGTACACCCCGGTTAAGCAACTGCGTGCCTGCTATTTGTAAGAGATCCCTGCC 380
QY 247 CTGGATCCCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 306
Db 381 CTGGATCCCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 440
QY 307 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 366
Db 441 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 500
QY 367 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAACTGTTTACCG 426
Db 501 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAACTGTTTACCG 560
QY 427 TTATGTGCGGATGTTGAAACCGTGGCTTGGCGGTGACCGTCTGTTGTTCAACGTTA 486
Db 561 TTATGTGCGGATGTTGAAACCGTGGCTTGGCGGTGACCGTCTGTTGTTCAACGTTA 620
QY 487 AGCGGCGCGAGTCTGCGTTACAAAATTTGGGTAAACCGCTTCTGTTAGTAACTGTTTACCG 666
Db 621 AGCGGCGCGAGTCTGCGTTACAAAATTTGGGTAAACCGCTTCTGTTAGTAACTGTTTACCG 800
QY 667 GCGTCACCGTTGTTACTAA 684
Db 801 GCGTCACCGTTGTTACTAA 818

RESULT 4
US-09-064-693A-25
Sequence 25, Application US/09064693A
Patent No. 6210937
GENERAL INFORMATION:
APPLICANT: Ward, Thomas E.
TITLE OF INVENTION: DEVELOPMENT OF GENETICALLY
ENGINEERED BACTERIA FOR PRODUCTION
OF A SPECIFIC PLASTICS PRECURSOR
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: W. Gary Goodson
ADDRESS: INEL--lockheed Martin Idaho
ADDRESSEE: Technologies Co.
STREET: P.O. Box 1625
CITY: Idaho Falls
STATE: Idaho
COUNTRY: USA
ZIP: 83415-3810
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: Toshiba Satellite Pro T2150CDS
OPERATING SYSTEM: Windows95

SOFTWARE: Word Perfect 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,693A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: W. Gary Goodson
REGISTRATION NUMBER: 22,387
REFERENCE/DOCKET NUMBER: LIT-PI-296
TELECOMMUNICATION INFORMATION:
TELEPHONE: (208)526-9469
TELEFAX: (208)526-8339
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 6641 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-064-693A-25
Query Match 72.8%; Score 498; DB 3; Length 6641;
Best Local Similarity 100.0%; Pred. No. 2.2e-164;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 ATGTACACCCCGGTTAAGCAACTGCGTGCCTGCTATTTGTAAGAGATCCCTGCC 246
Db 321 ATGTACACCCCGGTTAAGCAACTGCGTGCCTGCTATTTGTAAGAGATCCCTGCC 380
QY 247 CTGGATCCCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 306
Db 381 CTGGATCCCAACTGCTGCACTGGCTGTTGCTGGAGATTCCATGACAAAACGTTTGA 440
QY 307 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 366
Db 441 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGAGCAGAAATGA 500
QY 367 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAACTGTTTACCG 426
Db 501 ATCCCGGAAGAACTGCGCTGCGCGAAGAGTCTGTTACTGTTAGTAACTGTTTACCG 560
QY 427 TTATGTGCGGATGTTGAAACCGTGGCTTGGCGGTGACCGTCTGTTGTTCAACGTTA 486
Db 561 TTATGTGCGGATGTTGAAACCGTGGCTTGGCGGTGACCGTCTGTTGTTCAACGTTA 620
QY 487 AGCGGCGCGAGTCTGCGTTACAAAATTTGGGTAAACCGCTTCTGTTAGTAACTGTTTACCG 666
Db 621 AGCGGCGCGAGTCTGCGTTACAAAATTTGGGTAAACCGCTTCTGTTAGTAACTGTTTACCG 800
QY 667 GCGTCACCGTTGTTACTAA 684
Db 801 GCGTCACCGTTGTTACTAA 818

RESULT 5
US-09-896-866B-12
Sequence 12, Application US/09896866B
Patent No. 6642036
GENERAL INFORMATION:
APPLICANT: Flint, Dennis
APPLICANT: Meyer, Knut
APPLICANT: Viitanen, Paul
TITLE OF INVENTION: Sinapoylglucose:Malate Sinapoyltransferase Form Malate Conjugate

```
; TITLE OF INVENTION: Benzoin Acid Glucosides
; FILE REFERENCE: BC1034 US NA
; CURRENT APPLICATION NUMBER: US/09/896,866B
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/216,615
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-896-866B-12

Query Match      72.4%; Score 495; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 5.1e-164;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTCGGTGCGCTATTGTAAAGAGATCCCTGCC 246
Db 1 ATGTCACACCCCGCGTTAACGCAACTCGGTGCGCTATTGTAAAGAGATCCCTGCC 60

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTTGAA 306
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTTGAA 120

QY 307 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 366
Db 121 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 180

QY 367 ATCCCGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 426
Db 181 ATCCCGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 240

QY 427 TTATGTCCGATGTTGAACCGTGGCTGCTGCGGAAAGTCTCGTTACTGTTACGTGAAATTTTG 486
Db 241 TTATGTCCGATGTTGAACCGTGGCTGCTGCGGAAAGTCTCGTTACTGTTACGTGAAATTTTG 300

QY 487 AGCGGGCGGAGTGGCGTTACAAAATTGGGTAAACCGGTTAGGACGCTATCTGTTTC 546
Db 301 AGCGGGCGGAGTGGCGTTACAAAATTGGGTAAACCGGTTAGGACGCTATCTGTTTC 360

QY 547 ACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGTATGCGGGCTGTGGGG 606
Db 361 ACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGTATGCGGGCTGTGGGG 420

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGGTTAGGACGCTGTATGCGGGCTGTGGGG 666
Db 421 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGGTTAGGACGCTGTATGCGGGCTGTGGGG 480

QY 667 GCGTCACCGTTGTAC 681
Db 481 GCGTCACCGTTGTAC 495
```

RESULT 6

```
US-09-855-341-3
; Sequence 3, Application US/09855341
; Patent No. 6683231
; GENERAL INFORMATION:
; APPLICANT: VIITANEN, PAUL V.
; APPLICANT: MEYER, KNUT
; APPLICANT: VAN DYK, DREW
; TITLE OF INVENTION: HIGH LEVEL PRODUCTION OF P-HYDROXYBENZOIC ACID
; TITLE OF INVENTION: IN GREEN PLANTS
; FILE REFERENCE: BC1015 US NA
; CURRENT APPLICATION NUMBER: US/09/855,341
; CURRENT FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 3
; LENGTH: 495
; TYPE: DNA
```

```
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: E. coli
US-09-855-341-3

Query Match      72.4%; Score 495; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 5.1e-164;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTCGGTGCGCTATTGTAAAGAGATCCCTGCC 246
Db 1 ATGTCACACCCCGCGTTAACGCAACTCGGTGCGCTATTGTAAAGAGATCCCTGCC 60

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTTGAA 306
Db 61 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTTGAA 120

QY 307 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 366
Db 121 CAGCAGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 180

QY 367 ATCCCGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 426
Db 181 ATCCCGGGAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTTCGAGCAGATGAA 240

QY 427 TTATGTCCGATGTTGAACCGTGGCTGCTGCGGAAAGTCTCGTTACTGTTACGTGAAATTTTG 486
Db 241 TTATGTCCGATGTTGAACCGTGGCTGCTGCGGAAAGTCTCGTTACTGTTACGTGAAATTTTG 300

QY 487 AGCGGGCGGAGTGGCGTTACAAAATTGGGTAAACCGGTTAGGACGCTATCTGTTTC 546
Db 301 AGCGGGCGGAGTGGCGTTACAAAATTGGGTAAACCGGTTAGGACGCTATCTGTTTC 360

QY 547 ACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGTATGCGGGCTGTGGGG 606
Db 361 ACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCGCGTGTATGCGGGCTGTGGGG 420

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGGTTAGGACGCTGTATGCGGGCTGTGGGG 666
Db 421 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGGTTAGGACGCTGTATGCGGGCTGTGGGG 480

QY 667 GCGTCACCGTTGTAC 681
Db 481 GCGTCACCGTTGTAC 495

RESULT 7
US-09-489-039A-1234
; Sequence 1234, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1234
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1234

Query Match      41.2%; Score 282; DB 3; Length 645;
Best Local Similarity 72.9%; Pred. No. 1e-88;
Matches 363; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTCGGTGCGCTATTGTAAAGAGATCCCTGCC 246
Db 148 ATGTCACACCCCGCGTTAACGCAACTCGGTGCGCTATTGTAAAGAGATCCCTGCC 207
```


QY 247 CTGGATCCCAACTGCTGCATGGCTGTTGCTGGAGATTCCATGACAAAAAGTTTGA 306
Db |||||
208 CTCGCCGCCGCTAAGTGACTGGCTGCTGGAGACTCAATGACCCAGGCTTTGA 267
QY 307 CAGCAGGGAACCGTAGCGTACGATCCCGAAGGCTTTGTCGAGCAGATGAA 366
Db |||||
268 CAACAAGGAACGAGTCAACCTGACCTGGTTAACAAGGGTATATCGGCCGTGACGG 327
QY 367 ATCCCGGAGAACTGCGCTGCTGCGAAAGAGTCTCGTTACTGTTAGTGAAATTTTG 426
Db |||||
328 CTGACTGACGAGCGGCTGCTGCCGAGAGCGCGTACTGGCTGCGGAGATTATA 387
QY 427 TTATGTGCGGATGGTGAACCGTGGCTGTCGCGTACCGTCTGCTGTTCAACGTTA 486
Db |||||
388 CTAAATGCGGATGGGAGCCCTGGCTGCGCGGCGCAGCTGCGCGGAGTTCGACGCTG 447
QY 487 ACGGCGCGGAGTGGCTGTTACAAAATTTGGGTAAACCGCTTAGGAGCGTATCTGTTTC 546
Db |||||
448 TGTGTCGCGAGTGGCTGCTACAGAGCTCGGCGAACTCCGCTGGCGCGATACCTGTTT 507
QY 547 ACATCATGACATTAACCGGAGCTTTATTGAGATAGGCGCTGATGCCGCGCTGTGGGG 606
Db |||||
508 AGTCTGACATTAACCGCGATTTATTGAAATTTGGTTCGATGCGAGCGCTGTGGGG 567
QY 607 CGAGTTCGCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTTACG 666
Db |||||
568 CGTGTTCGCGCTGCGGCTGAGCGGCAACCCCTGCTGCTGACCGAGCTTTTTTTGCT 627
QY 667 GCGTCACGCTTACTAA 684
Db |||||
628 GCGTCGCGCTTACTAA 645

RESULT 8

US-09-839-477-3

; Sequence 3, Application US/09839477
; Patent No. 6723895
; GENERAL INFORMATION:
; APPLICANT: Shortish, Basil S.
; APPLICANT: DeBonte, Lorin R.
; TITLE OF INVENTION: PLANTS CONTAINING A CYTOSOLIC ACETYL
; COA-CARBOXYLASE
; FILE REFERENCE: 07148-094001
; CURRENT APPLICATION NUMBER: US/09/839,477
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/198,794
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(204)

US-09-839-477-3

Query Match 20.6%; Score 140.8; DB 3; Length 204;
Best Local Similarity 85.3%; Pred. No. 3.6e-39;
Matches 157; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGCTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACAAAGCT 60
Db |||||
1 ATGGCTTCCTCAGTTCTTTCTCTGAGGATGTCACCGCAGCAATGTTGCTCAAGCT 60
QY 61 AGCATGGTGCACCTTTCATCTGCTCAAAATCTTACGCACTTTCCTGTTACAAAGAG 120
Db |||||
61 AACATGGTGCACCTTTCATCTGCTCAAAATCTTACGCACTTTCCTGTTACAAAGAG 120
QY 121 CAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180
Db |||||

Db 121 CAAACCTTGACATCACTTCCATTTCCAGCAACGCGGAGAGTGCATGCAATGCAAGTG 180
QY 181 TGGC 184
Db |||||
181 TGGC 184

RESULT 9

US-09-495-797-41
; Sequence 41, Application US/09495797
; Patent No. 6369296
; GENERAL INFORMATION:
; APPLICANT: Ratcliff, Frank G
; APPLICANT: Martin-Hernandez, Ana M
; APPLICANT: Baulcombe, David C
; TITLE OF INVENTION: Viral Vectors
; FILE REFERENCE: Mewburn 43,047
; CURRENT APPLICATION NUMBER: US/09/495,797
; CURRENT FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N. benthamiana rubisco partial cDNA sequence

US-09-495-797-41

Query Match 20.6%; Score 140.8; DB 3; Length 499;
Best Local Similarity 85.3%; Pred. No. 6.3e-39;
Matches 157; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGCTCATTTCTTTCAGCAGCTGTTGCCACACGAGCAATGTTACAAAGCT 60
Db |||||
8 ATGGCTTCCTCAGTTCTTTCTCTGAGCAGTTGCCACCGCAGCAATGTTGCTCAAGCT 67
QY 61 AGCATGGTGCACCTTTCATCTGCTCAAAATCTTACGCACTTTCCTGTTACAAAGAG 120
Db |||||
68 AACATGGTGCACCTTTCATCTGCTCAAAATCTTACGCACTTTCCTGTTACAAAGAG 127
QY 121 CAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAAGTG 180
Db |||||
128 CAAACCTTGACATCACTTCCATTTGCCACCAACGCGGAGAGTGCATGCAAGTG 187
QY 181 TGGC 184
Db |||||
188 TGGC 191

RESULT 10

US-08-152-483B-8
; Sequence 8, Application US/08152483B
; Patent No. 5529909
; GENERAL INFORMATION:
; APPLICANT: Della-Cioppa, Guy
; APPLICANT: Kumagai, Monto
; TITLE OF INVENTION: TYROSINASE-ACTIVATOR
; TITLE OF INVENTION: PROTEIN FUSION ENZYME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 2730 Sand Hill Road
; CITY: Menlo Park
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0,
; SOFTWARE: Version #1.25


```
;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-331-004A-6

Query Match      19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACAGCAGCAATGTTCTCAAGCT 60

QY 61 AGCATGGTTGCACCTTTCACCTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
Db 61 AACATGGTGGCGCTTTTCACCTGCTTAAAGTCAGCTGCTCATTCCTGTTTCAAGGAAG 120

QY 121 CAAAACCTTGACATCACTTCCATTCCTAGTAGCATGTGGAGAGTAGTGCATGCAG 177
Db 121 CAAAACCTTGACATCACTTCCATTCCTAGTAGCATGTGGAGAGTAGTGCATGCAG 177

RESULT 15
US-08-096-623A-12
; Sequence 12, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Welsh & Katz, Ltd.
; STREET: 120 S. Riverside Plaza, 22nd Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,623A
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,061
; FILING DATE: 09-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/662,921
; FILING DATE: 28-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/562,674
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/525,551
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,613
; FILING DATE: 02-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: AMO-006.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 655-1500
; TELEFAX: (312) 655-1501
```

```
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; MAP POSITION: 1 to 177
; UNITS: bp
US-08-096-623A-12

Query Match      19.1%; Score 130.6; DB 2; Length 177;
Best Local Similarity 83.6%; Pred. No. 1.3e-35;
Matches 148; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACAGCAGCAATGTTACACAAGCT 60
Db 1 ATGGCTTCCTGTCATTTCTTTCAGCAGCTGTTGCCACAGCAGCAATGTTCTCAAGCT 60

QY 61 AGCATGGTTGCACCTTTCACCTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTTACAAAGAAG 120
Db 61 AACATGGTGGCGCTTTTCACCTGCTTAAAGTCAGCTGCTCATTCCTGTTTCAAGGAAG 120

QY 121 CAAAACCTTGACATCACTTCCATTCCTAGTAGCATGTGGAGAGTAGTGCATGCAG 177
Db 121 CAAAACCTTGACATCACTTCCATTCCTAGTAGCATGTGGAGAGTAGTGCATGCAG 177
```

Search completed: July 11, 2006, 19:21:21
Job time : 192 secs

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: July 11, 2006, 19:10:06 ; Search time 4407 Seconds
 (without alignments)
 8679.107 Million cell updates/sec

Title: US-10-718-311-7
 Perfect score: 684
 Sequence: 1 agggcttcctgtcatttc.....cggcgtcaccgtgtactaa 684

Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : EST:*

1: gb_est1:*
 2: gb_est3:*
 3: gb_est4:*
 4: gb_est5:*
 5: gb_est6:*
 6: gb_est7:*
 7: gb_est8:*
 8: gb_est9:*
 9: gb_est10:*
 10: gb_est11:*
 11: gb_est12:*
 12: gb_est13:*
 13: gb_est14:*
 14: gb_est15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	363	53.1	832	13	CL666350
2	295.4	43.2	828	13	CL687382
3	184	26.9	205	7	BE462847
4	184	26.9	235	7	AW039464
5	184	26.9	237	7	AW039041
6	184	26.9	238	7	AW041024
7	184	26.9	250	7	AW442951
8	184	26.9	260	7	AW037699
9	184	26.9	260	7	AW094122
10	184	26.9	267	1	AI773928
11	184	26.9	277	1	AI773502
12	184	26.9	279	2	BI931188
13	184	26.9	279	7	BE462486
14	184	26.9	283	7	AW094091
15	184	26.9	286	7	BF051382
16	184	26.9	288	7	AW037811
17	184	26.9	294	7	AW040243
18	184	26.9	329	7	AW038470
19	184	26.9	330	7	AW092013

20	184	26.9	351	2	BG129648	BG129648	EST475294
21	184	26.9	356	2	BI927635	BI927635	EST547524
22	184	26.9	367	7	AW039845	AW039845	EST282318
23	184	26.9	370	1	AI773548	AI773548	EST254648
24	184	26.9	372	7	BE463004	BE463004	EST252666
25	184	26.9	373	7	AW093805	AW093805	EST286985
26	184	26.9	375	2	BG123527	BG123527	EST489173
27	184	26.9	377	7	AW038793	AW038793	EST280749
28	184	26.9	382	1	AI777051	AI777051	EST252018
29	184	26.9	385	7	AW092154	AW092154	EST285250
30	184	26.9	393	7	AW040522	AW040522	EST283482
31	184	26.9	394	7	BG128649	BG128649	EST474295
32	184	26.9	395	7	AW217063	AW217063	EST295777
33	184	26.9	401	7	AW093581	AW093581	EST286761
34	184	26.9	404	2	BM535307	BM535307	EST588329
35	184	26.9	407	2	BI929701	BI929701	EST549590
36	184	26.9	408	7	AW038880	AW038880	EST280836
37	184	26.9	412	2	BI928909	BI928909	EST548798
38	184	26.9	416	7	AW094716	AW094716	EST287896
39	184	26.9	417	7	AW933080	AW933080	EST3358923
40	184	26.9	420	7	AW443412	AW443412	EST308342
41	184	26.9	420	7	BG133589	BG133589	EST466577
42	184	26.9	421	2	BG642684	BG642684	EST510878
43	184	26.9	421	2	BG642684	BG642684	EST510878
44	184	26.9	422	1	AI775340	AI775340	EST256440
45	184	26.9	423	7	AW041189	AW041189	EST284053

ALIGNMENTS

RESULT 1
 CL666350/c

LOCUS
 DEFINITION

CL666350

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CL666350 832 bp DNA linear GSS 09-JUL-2004
 PRI0152b A03 - PRI0152b.B21 (832) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 CL666350 GI:50159025
 GSS.
 Pristionchus pacificus
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 832)
 Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
 AppaDB: an AcedB database for the nematode satellite organism
 Pristionchus pacificus
 Nucleic Acids Res. 32 (1), D421-D422 (2004)
 14681447
 Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: raif.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: T7
 Class: fosmid ends.
 Location/Qualifiers
 1..832
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 53.1%; Score 363; DB 13; Length 832;

Best Local Similarity 100.0%; Pred. No. 1.8e-104; Mismatches 0; Indels 0; Gaps 0; Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 GTAAGCGTGACGATGATCCGGAAGGGTTTGTGAGCAGAGATGAATCCCGAAGAACTG 381
 |||||
 Db 832 GTAAGCGTGACGATGATCCGGAAGGGTTTGTGAGCAGAGATGAATCCCGAAGAACTG 773

QY 382 CGCTGCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAAATTTTGTATGTGCCGATGGT 441
 |||||
 Db 772 CGCTGCTGCCGAAAGAGTCTCGTTACTGTTTACGTGAAATTTTGTATGTGCCGATGGT 713

QY 442 GAACCGTGGCTTGGCCGGTGTACCGTCTGTTCTGTCAACGTTAAGCGGGCCGGAGCTG 501
 |||||
 Db 712 GAACCGTGGCTTGGCCGGTGTACCGTCTGTTCTGTCAACGTTAAGCGGGCCGGAGCTG 653

QY 502 GGGTTACAAAATTGGGTAAACGCCGTTAGACGCTATCTGTTACATCATCGACATTA 561
 |||||
 Db 652 GGGTTACAAAATTGGGTAAACGCCGTTAGACGCTATCTGTTACATCATCGACATTA 593

QY 562 ACCCGGGACTTTATTGAGATAGCCGCTGATGCGGGCTGTGGGGCGGACGTTCCCGCCTG 621
 |||||
 Db 592 ACCCGGGACTTTATTGAGATAGCCGCTGATGCGGGCTGTGGGGCGGACGTTCCCGCCTG 533

QY 622 CGATTAAAGCGGTAACCGCTGTGTGCTAAACAGAACTGTTTTTACC GGCGTCAACGTTGTAC 681
 |||||
 Db 532 CGATTAAAGCGGTAACCGCTGTGTGCTAAACAGAACTGTTTTTACC GGCGTCAACGTTGTAC 473

QY 682 TAA 684
 |||||
 Db 472 TAA 470

RESULT 2
 CL687382
 LOCUS
 DEFINITION
 Pristionchus pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Pristionchus pacificus
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.

REFERENCE
 1 (bases 1 to 828)
 Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.
 AppADB: an AcedB database for the nematode satellite organism
 Pristionchus pacificus
 Nucleic Acids Res. 32 (1), D421-D422 (2004)

JOURNAL
 PUBMED
 COMMENT
 Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel.: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: 17
 Class: fosmid ends.

FEATURES
 source
 Location/Qualifiers
 1..828
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus var. California"
 /note="Vector: pEpifos-5 Fosmid vector"

Query Match 43.2%; Score 295.4; DB 13; Length 828; Best Local Similarity 99.7%; Pred. No. 8.7e-83; Mismatches 296; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 187 ATGTACACACCCCGGTTAAACGCAACTGCGTGGCTGCGCTATTGTAAAGAGATCCCTGCC 246
 |||||
 Db 532 ATGTACACACCCCGGTTAAACGCAACTGCGTGGCTGCGCTATTGTAAAGAGATCCCTGCC 591

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGTTTGA 306
 |||||
 Db 592 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAAAGTTTGA 651

QY 307 CAGCAGGAAAAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTCCAGCAGAAATGA 366
 |||||
 Db 652 CAGCAGGAAAAACCGTAAAGCGTGACGATGATCCGCGAAGGTTTGTCCAGCAGAAATGA 711

QY 367 ATCCCGGAAAGTCCGCTGCTGCGGAAAGAGTCTCGTTACTGTTACGTGAAATTTTG 426
 |||||
 Db 712 ATCCCGGAAAGTCCGCTGCTGCGGAAAGAGTCTCGTTACTGTTACGTGAAATTTTG 771

QY 427 TTATGTGCCGATGTGAACCGTGGCTTCCCGTGTACCGTCTGTTCTGTGTCAACG 483
 |||||
 Db 772 TTATGTGCCGATGTGAACCGTGGCTTCCCGTGTACCGTCTGTTCTGTGTCAACG 828

RESULT 3
 BE462847
 LOCUS
 DEFINITION
 EST325226 tomato flower buds 0-3 mm, mRNA linear EST 18-MAY-2001
 Lycopersicon esculentum cDNA clone cFOA16011, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (Solanum lycopersicum)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
 1 (bases 1 to 205)
 van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E., Liang, P., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M., Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
 Generation of ESTs from tomato flower tissue, 0-3 mm buds
 Contact: CUGI
 Unpublished (1999)

TITLE
 JOURNAL
 COMMENT
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 Location/Qualifiers
 1..205
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cFOA16011"
 /tissue_type="flower"
 /dev_stage="0-3mm buds"
 /clone_lib="tomato flower buds 0-3 mm, Cornell University"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tankley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

ORIGIN
 Query Match 26.9%; Score 184; DB 7; Length 205; Best Local Similarity 100.0%; Pred. No. 2.9e-47; Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTGTGTCATTCTTCAGCAGCTGTTCACGACGCAATGTTACAAAGCT 60

```
|||||
8 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 67
|||||
61 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
|||||
68 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 127
|||||
121 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
128 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 187
|||||
181 TGGC 184
188 TGGC 191
|||||

RESULT 4
AW039464
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 235)
AUTHORS
D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,
Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
Giovannoni, J.
TITLE
Generation of ESTs from tomato leaf tissue
JOURNAL
Unpublished (1999)
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
1..235
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET10112"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="Xl1-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 235;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 60
DB 14 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 73
QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
DB 74 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 133
|||||
```

```
121 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
134 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 193
|||||
181 TGGC 184
194 TGGC 197
|||||

RESULT 5
AW039041
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (Solanum lycopersicum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE
1 (bases 1 to 237)
AUTHORS
D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,
Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
Giovannoni, J.
TITLE
Generation of ESTs from tomato leaf tissue
JOURNAL
Unpublished (1999)
COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
source
1..237
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rio Grande Ptor"
/db_xref="taxon:4081"
/clone="cLET10C2"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="Xl1-Blue MRF"
/clone_lib="tomato mixed elicitor, BTI"
/notes="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLET - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 237;
Best Local Similarity 100.0%; Pred. No. 3e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 60
DB 9 ATGGCTTCTCTGTCAATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAAGCT 68
QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
DB 69 AGCATGGTTGCACCTTTCACTGGTCTCAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 128
QY 121 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 180
DB 129 CAAAACCTTGACATCACTTCCATGCTAGCAATGGTGGAGAGTTAGCTGCATGCAGGTG 188
|||||
181 TGGC 184
189 TGGC 192
|||||
```


asteroids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 260)
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.
 Generation of ESTs from tomato leaf tissue
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..260
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande PtoR"
 /db_xref="taxon:4081"
 /clone="cLET4M17"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

ORIGIN
 Query Match 26.9%; Score 184; DB 7; Length 260;
 Best Local Similarity 100.0%; Pred. No. 3.1e-47;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAGCT 60
 12 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAGCT 71
 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
 72 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 131
 121 CAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAGAGTTAGTGCATGCGAGGTG 180
 132 CAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAGAGTTAGTGCATGCGAGGTG 191
 181 TGGC 184
 192 TGGC 195

RESULT 9
 AW094122
 LOCUS
 DEFINITION
 EST287302 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 clone cLET27G20, mRNA sequence.
 ACCESSION
 AW094122
 VERSION
 AW094122.1 GI:6059717
 KEYWORDS
 EST.
 SOURCE
 Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 260)
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
 Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.
 Generation of ESTs from tomato leaf tissue

Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 Location/Qualifiers
 1..260
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="Rio Grande PtoR"
 /db_xref="taxon:4081"
 /clone="cLET27G20"
 /tissue_type="leaf"
 /dev_stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /clone_lib="tomato mixed elicitor, BTI"
 /notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
 okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."

ORIGIN
 Query Match 26.9%; Score 184; DB 7; Length 260;
 Best Local Similarity 100.0%; Pred. No. 3.1e-47;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAGCT 60
 6 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTGGCCACGCGAGCAATGTTACACAGCT 65
 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
 66 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 125
 121 CAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAGAGTTAGTGCATGCGAGGTG 180
 126 CAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAGAGTTAGTGCATGCGAGGTG 185
 181 TGGC 184
 186 TGGC 189

RESULT 10
 AW094122
 LOCUS
 DEFINITION
 EST255028 tomato resistant, Cornell Lycopersicon esculentum cDNA
 clone cLER8L7, mRNA sequence.
 ACCESSION
 AW094122
 VERSION
 AW094122.1 GI:5271969
 KEYWORDS
 EST.
 SOURCE
 Lycopersicon esculentum (Solanum lycopersicum)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 267)
 D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,
 Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
 Giovannoni, J.J. and Martin, G.B.
 Generation of ESTs from Pseudomonas resistant tomato
 Unpublished (1999)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

```

FEATURES
  source
    Location/Qualifiers
      1..267
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /cultivar="R11-12 (35S)::Pto in Rio Grande x Money Maker)"
        /db_xref="taxon:4081"
        /clone="CLER817"
        /tissue_type="leaf"
        /dev_stage="4-week old"
        /lab_host="SOLR"
        /clone_lib="tomato resistant, Cornell"
        /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
  Query Match      26.9%; Score 184; DB 1; Length 267;
  Best Local Similarity 100.0%; Pred. No. 3.2e-47;
  Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTGTGTCATTTCTTCCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 60
    |||
Db 23 ATGGCTTCTGTGTCATTTCTTCCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 82
    |||

QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAAACTTTCAGCCACTTTCCCTGTTACAAAGAAG 120
    |||
Db 83 AGCATGGTTGCACCTTTCAGTGGTCTCAAACTTTCAGCCACTTTCCCTGTTACAAAGAAG 142
    |||

QY 121 CAAACCTTGACATCATTCCATTCAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
    |||
Db 143 CAAACCTTGACATCATTCCATTCAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 202
    |||

QY 181 TGGC 184
    |||
Db 203 TGGC 206

RESULT 11
BI773502
LOCUS BI773502 277 bp mRNA linear EST 18-MAY-2001
DEFINITION EST254602 tomato resistant, Cornell Lycopersicon esculentum cDNA
clone CLER7A10, mRNA sequence.
ACCESSION AI773502
VERSION AI773502.1 GI:5271543
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 277)
AUTHORS D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
Giovannoni,J.J. and Martin,G.B.
Generation of ESTs from Pseudomonas resistant tomato
TITLE Lycopersicon esculentum (Solanum lycopersicum)
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES
  source
    Location/Qualifiers
      1..277
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /cultivar="R11-12 (35S)::Pto in Rio Grande x Money Maker)"
        /db_xref="taxon:4081"
        /clone="CLER7A10"
        /tissue_type="leaf"
        /dev_stage="4-week old"

/clone_lib="tomato resistant, Cornell"
/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; cLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
  Query Match      26.9%; Score 184; DB 1; Length 277;
  Best Local Similarity 100.0%; Pred. No. 3.2e-47;
  Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCTGTGTCATTTCTTCCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 60
    |||
Db 5 ATGGCTTCTGTGTCATTTCTTCCAGCAGCTGTGGCCACGAGCAATGTTACACAAGCT 64
    |||

QY 61 AGCATGGTTGCACCTTTCAGTGGTCTCAAACTTTCAGCCACTTTCCCTGTTACAAAGAAG 120
    |||
Db 65 AGCATGGTTGCACCTTTCAGTGGTCTCAAACTTTCAGCCACTTTCCCTGTTACAAAGAAG 124
    |||

QY 121 CAAACCTTGACATCATTCCATTCAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
    |||
Db 125 CAAACCTTGACATCATTCCATTCAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 184
    |||

QY 181 TGGC 184
    |||
Db 185 TGGC 188

RESULT 12
BI931188
LOCUS BI931188 279 bp mRNA linear EST 18-OCT-2001
DEFINITION EST551077 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cT0C19C6 5' end, mRNA sequence.
ACCESSION BI931188
VERSION BI931188.1 GI:16245660
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 279)
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
TITLE Lycopersicon esculentum (Solanum lycopersicum)
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.

FEATURES
  source
    Location/Qualifiers
      1..279
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="cT0C19C6"
        /tissue_type="flower"
        /dev_stage="buds 8mm to preanthesis"
        /clone_lib="tomato flower, 8 mm to preanthesis buds"
        /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

ORIGIN

Query Match 26.9%; Score 184; DB 2; Length 279;
 Best Local Similarity 100.0%; Pred. No. 3.2e-47;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 60
 DB 7 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 66

QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
 DB 67 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 126

QY 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
 DB 127 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 186

QY 181 TGGC 184
 DB 187 TGGC 190

RESULT 13

BE462486
 LOCUS EST324752 tomato flower buds 0-3 mm, Cornell University
 DEFINITION Lycopersicon esculentum cDNA clone cTOAL301, mRNA sequence.

ACCESSION

BE462486

VERSION

BE462486.1 GI:9508255

KEYWORDS

EST.

SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 279)

van der Hoeven, R.S., Bezzerides, J.L., Matern, A.L., Holt, I.E.,

Liang, F., Hansen, T., Craven, M.B., Bowman, C.L., Ronning, C.M.,

Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and

Tanksley, S.D.

Generation of ESTs from tomato flower tissue, 0-3 mm buds

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

Location/Qualifiers

1..279

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="cTOAL301"

/tissue_type="flower"

/dev_stage="0-3mm buds"

/clone_lib="tomato flower buds 0-3 mm, Cornell University"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; supplier: Tanksley; Flower buds and flowers were

taken from greenhouse plants (4-8 wks old, TA496). They

were immediately frozen in liquid nitrogen and then

size-separated while remaining frozen."

ORIGIN

Query Match 26.9%; Score 184; DB 7; Length 279;
 Best Local Similarity 100.0%; Pred. No. 3.2e-47;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 60
 DB 7 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 66

QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
 DB 67 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 126

QY 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180
 DB 127 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 186

QY 181 TGGC 184
 DB 187 TGGC 190

RESULT 14

AW094091
 LOCUS EST287271 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION clone cLET27A10, mRNA sequence.

ACCESSION

AW094091

VERSION

AW094091.1 GI:6059686

KEYWORDS

EST.

SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 283)

D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,

Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W.,

Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and

Giovannoni, J.J.

Generation of ESTs from tomato leaf tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

Location/Qualifiers

1..283

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="Rio Grande PtoR"

/db_xref="taxon:4081"

/clone="cLET27A10"

/tissue_type="leaf"

/dev_stage="4-6 week old plants"

/lab_host="XLI-Blue MRF"

/clone_lib="tomato mixed elicitor, BTI"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; cLET - Inoculated with a variety of disease response

elicitors. Plants exposed to 2,6 dichloroisonicotinic

acid, BTH, jasmonic acid, ethylene, fenthion, EIX,

okadaic acid, or systemin prior to tissue harvest. EcoRI

site was destroyed during cloning."

ORIGIN

Query Match 26.9%; Score 184; DB 7; Length 283;
 Best Local Similarity 100.0%; Pred. No. 3.2e-47;
 Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 60
 DB 12 ATGGCTTCCTCTGTCATTTCTTCAGCAGCTGTTCACACGAGCAATGTTACACAGCT 71

QY 61 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
 DB 72 AGCATGGTTGCACCTTTCACTGGTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 131

QY 121 CAAAACCTTGACATCACTTCCATTTGCTAGCAATGTTGGAAGAGTTAGCTGCATGCAGGTG 180

Db 132 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 191
QY 181 TGGC 184
|||
Db 192 TGGC 195

RESULT 15
BF051382
LOCUS BF051382 286 bp mRNA linear EST 18-MAY-2001
DEFINITION EST436557 tomato developing/immature green fruit Lycopersicon
esculentum cDNA clone cLEM22P19 5' sequence, mRNA sequence.
ACCESSION BF051382
VERSION BF051382.1 GI:10805278
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 286)
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksey,S.D.
REFERENCE Generation of ESTs from tomato fruit tissue, immature green
AUTHORS Unpublished (2000)
TITLE Contact: CUGI
JOURNAL Clemson University Genomics Institute
COMMENT Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
Location/Qualifiers

FEATURES
source
1..286
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEM22P19"
/tissue_type="fruit"
/dev_stage="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/clone_lib="tomato developing/immature green fruit"
/note="Vector: pBluescriptSKmCUadapt; Site 1: EcoRI;
Site 2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."

ORIGIN
Query Match 26.9%; Score 184; DB 7; Length 286;
Best Local Similarity 100.0%; Pred. No. 3.3e-47;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCCTCTGTGTCATTCTTTCAGCAGCTGTGTCACACGAGCAATGTTACACAAGCT 60
Db |||
10 ATGGCTTCCTCTGTGTCATTCTTTCAGCAGCTGTGTCACACGAGCAATGTTACACAAGCT 69
QY 61 AGCATGGTTGCACCTTCACCTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 120
Db |||
70 AGCATGGTTGCACCTTCACCTGGTCTCAAACTTCAGCCACTTCCCTGTTACAAAGAAG 129
QY 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 180
Db |||
130 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAGGTG 189
QY 181 TGGC 184
|||
Db 190 TGGC 193

Search completed: July 11, 2006, 21:36:48
Job time : 4410 secs

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 11, 2006, 18:50:32 ; Search time 588 Seconds
 (without alignments)
 8110.575 Million cell updates/sec

Title: US-10-718-311-7
 Perfect score: 684
 Sequence: 1 atggctctctctgtcatttc.....cggcgctaccgtgtgactaa 684

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 104899840

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : N Geneseq 8.*
 1: Geneseqn1980s.*
 2: Geneseqn1990s.*
 3: Geneseqn2000s.*
 4: Geneseqn2001as.*
 5: Geneseqn2001bs.*
 6: Geneseqn2002as.*
 7: Geneseqn2002bs.*
 8: Geneseqn2003as.*
 9: Geneseqn2003bs.*
 10: Geneseqn2003cs.*
 11: Geneseqn2003ds.*
 12: Geneseqn2004as.*
 13: Geneseqn2004bs.*
 14: Geneseqn2005s.*
 15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	684	100.0	684	6	ABA91838 Chloropla
2	684	100.0	684	6	ABK47712 DNA seque
3	684	100.0	684	10	ADC25988 Tomato/Es
4	684	100.0	684	14	ADV91622 Tomato ru
5	501	73.2	3452	14	ADV91636 C. paraps
6	499.4	73.0	1971	14	ADV91643 C. paraps
7	498	72.8	498	14	ADV91618 Escherich
8	498	72.8	854	4	AAD06956 ubiC gene
9	498	72.8	6641	4	AAD06957 pME2 comp
10	496.4	72.6	2000	2	AAT29820 E. coli u
11	495	72.4	495	2	AAQ92409 E. coli ch
12	495	72.4	495	6	ABA91837 Escherich
13	495	72.4	495	6	ABK47709 DNA seque
14	495	72.4	495	10	ADC25984 Escherich
15	472.8	69.1	1207	5	AA90335 DNA encod
16	472.8	69.1	1918	2	AAQ67671 E.coli ub
17	282	41.2	645	11	ACH95439 Klebsiell
18	145.6	21.3	492	10	ADC76164 DNA homol

19	144	21.1	581	10	ADC75089	N bentham
20	144	21.1	614	10	ADC76956	DNA homol
21	144	21.1	632	10	ADC76953	DNA homol
22	144	21.1	718	10	ADK58382	Plant DNA
23	144	21.1	736	10	ADK54321	Plant DNA
24	144	21.1	736	10	ADK57660	Plant DNA
25	144	21.1	841	11	ADM44879	Insect re
26	144	21.1	847	11	ADM45447	Insect re
27	144	21.1	924	10	ADC75566	DNA homol
28	144	21.1	958	10	ADC76165	DNA homol
29	142.4	20.8	297	10	ADK59826	Plant DNA
30	142.4	20.8	377	10	ADK56121	Plant DNA
31	142.4	20.8	489	10	ADK56120	Plant DNA
32	142.4	20.8	504	10	ADK59824	Plant DNA
33	142.4	20.8	608	10	ADC76944	DNA homol
34	142.4	20.8	615	10	ADC76949	DNA homol
35	142.4	20.8	684	10	ADC76948	DNA homol
36	142.4	20.8	859	11	ADM45065	Insect re
37	141.4	20.7	619	11	ADM45158	Insect re
38	140.8	20.6	204	6	AAI70688	Tobacco R
39	140.8	20.6	499	6	AAI35635	Nicotiana
40	140.8	20.6	609	10	ADK76965	DNA homol
41	137.6	20.1	668	10	ADK59825	Plant DNA
42	137.6	20.1	668	11	ADM45687	Insect re
43	133.2	19.5	1442	2	AAQ92327	Chloropla
44	132.6	19.4	171	14	ABE04209	Tomato rb
45	132.6	19.4	171	14	ABE04212	Tomato rb

ALIGNMENTS

RESULT 1

ABA91838
 ID ABA91838 standard; DNA; 684 BP.

AC ABA91838;
 XX

XX 29-AUG-2003 (revised)
 DT 15-MAY-2002 (first entry)

XX Chloroplast transit peptide-chorismate pyruvate lyase fusion gene.

XX Chloroplast transit peptide; tomato; Rubisco; plant;

KW ribulose-1,5-bisphosphate carboxylase; chorismate pyruvate lyase; CPL;

KW enzyme; p-hydroxybenzoic acid; transgenic plant; gene; ds.

XX Lycopersicon esculentum.

OS Escherichia coli.

XX Chimeric.

Key Location/Qualifiers

CDS 1..495

FT /*tag= a

FT /product= "chloroplast-targeted CPL fusion"

FT /*tag= b

FT /note= "tomato Rubisco chloroplast targeting sequence"

FT mat_peptide 187..492

FT /*tag= c

FT /note= "E. coli chorismate pyruvate lyase"

XX WO200194607-A2.

PN 13-DEC-2001.

XX 22-MAY-2001; 2001WO-US016661.

PD 02-JUN-2000; 2000US-0209854P.

XX (DUFO) DU PONT DE NEMOURS & CO E I.

XX Meyer K, Van Dyk DE, Viitanen PV;

PI

```
XX WPI; 2002-226795/28.
DR P-PSDB; AAM50959.
XX
XX Producing para-hydroxy benzoic acid in green plant, comprises expression
PT of unique expression cassette containing gene encoding chorismate
PT pyruvate lyase operably linked to specific chloroplast targeting
PT sequence.
XX
XX Claim 17; Page 57; 60pp; English.
XX
XX The present nucleotide sequence is that of an open reading frame encoding
CC chloroplast-targeted chorismate pyruvate lyase (CPL) fusion protein TP-
CC CPL (see AAM50959). PCR (see ABA91841-42) was used to generate a DNA
CC fragment corresponding to the transit peptide of tomato Rubisco small
CC subunit and the first 4 amino acid residues of mature Rubisco. The DNA
CC fragment was ligated into pET24a-CPL, which carries the Escherichia coli
CC CPL open reading frame (see ABA91837). The TP-CPL construct is an example
CC of expression cassettes of the invention that are designed for the high-
CC level production of p-hydroxybenzoic acid (pHBA) in higher plants.
CC Cleavage of TP-CPL in the chloroplast releases a novel polypeptide (see
CC AAM50961) that has full enzyme activity, converting chorismate to pHBA. A
CC claimed plant comprising a CPL expression cassette is selected from
CC soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
CC barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,
CC millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric
CC component of liquid crystal polymers which have application in the
CC automotive, electrical and other industries. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
XX Query Match 100.0%; Score 684; DB 6; Length 684;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-219; Indels 0; Gaps 0;
XX Matches 684; Conservative 0; Mismatches 0;
XX
XX 1 ATGGCTTCTCTGTGATTTCTTTCAGCAGCTGTGGCCACACGAGCAATGTTACACAAGCT 60
DB |||||||
XX 1 ATGGCTTCTCTGTGATTTCTTTCAGCAGCTGTGGCCACACGAGCAATGTTACACAAGCT 60
DB |||||||
XX 61 AGCATGGTTGACCTTTTCAGTGGTCTCAAACTTCAGGCACCTTTCCCTGTTACAAAGAAG 120
DB |||||||
XX 61 AGCATGGTTGACCTTTTCAGTGGTCTCAAACTTCAGGCACCTTTCCCTGTTACAAAGAAG 120
DB |||||||
XX 121 CAAACACCTTGACATCAGTCTTCCATTCCTAGCAATGTTGGAAGTGTAGCTGCATGAGGTG 180
DB |||||||
XX 121 CAAACACCTTGACATCAGTCTTCCATTCCTAGCAATGTTGGAAGTGTAGCTGCATGAGGTG 180
DB |||||||
XX 181 TGGCATATGTACACCCCGCGTTAAACGCAACTGCGTGCCTGCTATTTGTAAGAGATC 240
DB |||||||
XX 181 TGGCATATGTACACCCCGCGTTAAACGCAACTGCGTGCCTGCTATTTGTAAGAGATC 240
DB |||||||
XX 241 CCTGCCCTGGATCCGCAACTGCTCAGTGGTGTGCTGGAGGATTCATGACAAAAGCT 300
DB |||||||
XX 241 CCTGCCCTGGATCCGCAACTGCTCAGTGGTGTGCTGGAGGATTCATGACAAAAGCT 300
DB |||||||
XX 301 TTTGAACAGCAGGAGAAACGGTAAAGCTGACATGATCCGCGAAGGGTTTGTCCAGCAG 360
DB |||||||
XX 301 TTTGAACAGCAGGAGAAACGGTAAAGCTGACATGATCCGCGAAGGGTTTGTCCAGCAG 360
DB |||||||
XX 361 AATGAAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
DB |||||||
XX 361 AATGAAATCCCGAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
DB |||||||
XX 421 ATTTTGTATGTGCGGATGGTGAACCGTGGCTTTCGGGTGCTGACCGTCTCTCTGTGCA 480
DB |||||||
XX 421 ATTTTGTATGTGCGGATGGTGAACCGTGGCTTTCGGGTGCTGACCGTCTCTCTGTGCA 480
DB |||||||
XX 481 ACGTTTAAGCGGCGCGAGCTGCGGTTACAAAATTTGGGTAAACCCCGTTAGGAGCTAT 540
DB |||||||
XX 481 ACGTTTAAGCGGCGCGAGCTGCGGTTACAAAATTTGGGTAAACCCCGTTAGGAGCTAT 540
DB |||||||
XX 541 CTGTTTCATCATCATGACATTAACCCGGGACTTTATTGAGATAGGCCGTGATGCCGGGCTG 600
DB |||||||
XX 541 CTGTTTCATCATCATGACATTAACCCGGGACTTTATTGAGATAGGCCGTGATGCCGGGCTG 600
DB |||||||
XX 601 TGGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
DB |||||||
XX 601 TGGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
DB |||||||
XX 661 TTACCGGCGGCACCGTTGTACTAA 684
DB |||||||
XX 661 TTACCGGCGGCACCGTTGTACTAA 684
DB |||||||
XX
XX RESULT 2
XX ABK47712
XX ID ABK47712 standard; DNA; 684 BP.
XX AC ABK47712;
XX DT 18-JUN-2002 (first entry)
XX DE DNA sequence of ORF for TP-CPL fusion protein.
XX KW SNG1; sinapoylglucose accumulator 1; SMT; aromatic ester; plant;
XX sinapoylglucose:malate sinapoyltransferase; glycosylated aromatic acid;
XX malate conjugated aromatic acid; polymer synthesis; ubiC; TP-CPL; gene;
XX carboxylic acid conjugated aromatic acid; chorismate pyruvate lyase;
XX tomato Rubisco small subunit precursor; rbcS2; mutant; ds.
XX OS Escherichia coli.
XX LY Lycopersicon esculentum.
XX CH Chimeric.
XX
XX Key Location/Qualifiers
XX CDS 1..684
XX /*tag= a
XX /product= "TP-CPL fusion protein"
XX
XX WO200204653-A2.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US021283.
XX
XX 07-JUL-2000; 2000US-0216615P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Flint D, Meyer K, Viitanen PV;
XX
XX WPI; 2002-303779/34.
XX
XX P-PSDB; AAU77942.
XX
XX Producing aromatic acid conjugates, involves contacting glycosylated
PT aromatic acid with malate, alpha-hydroxycarboxylic acid or alcohol in the
PT presence of sinapoylglucose:malate sinapoyltransferase.
XX
XX Example; Page 70; 72pp; English.
XX
XX The present invention relates to the isolation of Arabidopsis thaliana
CC gene (SNG1, sinapoylglucose accumulator 1) encoding
CC sinapoylglucose:malate sinapoyltransferase (SMT). SMT catalyses the
CC substitution of a glucose moiety on a glycosylated aromatic acid with a
CC malate moiety to form a malate conjugated aromatic acid. The enzyme is
CC useful for producing malate conjugated aromatic acids, carboxylic acid
CC conjugated aromatic acids or aromatic esters. The malate conjugated
CC aromatic acids are useful in the synthesis of various polymers. The
CC present sequence representing the ORF (open reading frame) for TP-CPL
CC (tomato Rubisco small subunit precursor for rbcS2-Escherichia coli
CC chorismate pyruvate lyase (CPL)) fusion protein is described in the
CC examples of the present invention
XX
XX Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;
```

```
Query Match      100.0%; Score 684; DB 6; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.4e-219;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGCTTCTCTGCTGCTATTTCTTTCAGCAGCTGTTGCCACACGCAAGCAATGTTACACAAGCT 60
DB      1 ATGGCTTCTCTGCTGCTATTTCTTTCAGCAGCTGTTGCCACACGCAAGCAATGTTACACAAGCT 60

QY      61 AGCATGGTTGCACCTTTCTACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
DB      61 AGCATGGTTGCACCTTTCTACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

QY      121 CAAACCTTGACATCACTTCCATTTCTAGCAATGTTGGAGAGTTAGCTGCATGCAAGTG 180
DB      121 CAAACCTTGACATCACTTCCATTTCTAGCAATGTTGGAGAGTTAGCTGCATGCAAGTG 180

QY      181 TGGCATATGTCAACCCCGGTTAAACCAACTGCGTGCCTGCTATTGTAAGAAGATC 240
DB      181 TGGCATATGTCAACCCCGGTTAAACCAACTGCGTGCCTGCTATTGTAAGAAGATC 240

QY      241 CTGCGCTCGATCCGCAACTGCTGACTGCTGTTCTCGAGGATTTCCATGACAAACGT 300
DB      241 CTGCGCTCGATCCGCAACTGCTGACTGCTGTTCTCGAGGATTTCCATGACAAACGT 300

QY      301 TTTGAACACAGGAGAAACCGTGAACGATGATCCGCGAAGGTTTTCGAGCAG 360
DB      301 TTTGAACACAGGAGAAACCGTGAACGATGATCCGCGAAGGTTTTCGAGCAG 360

QY      361 AATGAAATCCCGAAGAACTGCGCGTGTGCGCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
DB      361 AATGAAATCCCGAAGAACTGCGCGTGTGCGCGAAAGAGTCTCGTTACTGGTTACGTGAA 420

QY      421 ATTTGTTATGTCGCGATGCTGAAACCGTGGCTTGCCTGCGTGCATCGCTGCTGTTGTC 480
DB      421 ATTTGTTATGTCGCGATGCTGAAACCGTGGCTTGCCTGCGTGCATCGCTGCTGTTGTC 480

QY      481 AGCTTAAGCGCGCGAGCTGGCTTACAAAATTTGGGTAAACCGCGTTAGACGCTAT 540
DB      481 AGCTTAAGCGCGCGAGCTGGCTTACAAAATTTGGGTAAACCGCGTTAGACGCTAT 540

QY      541 CTGTTACATATCGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTG 600
DB      541 CTGTTACATATCGACATTAACCCGGGACTTTATTGAGATAGGCGGTGATGCGGGCTG 600

QY      601 TGGGGCGACGTTCCCGCTGCGATTAACCGGTAAACCGCTGTTGCTAAACAGAACTGTTT 660
DB      601 TGGGGCGACGTTCCCGCTGCGATTAACCGGTAAACCGCTGTTGCTAAACAGAACTGTTT 660

QY      661 TTACCGGCGTACCGTTGTTACTAA 684
DB      661 TTACCGGCGTACCGTTGTTACTAA 684
```

```
RESULT 3
ADC25988
ID   ADC25988 standard; DNA; 684 BP.
XX
AC   ADC25988;
XX
DT   18-DEC-2003 (first entry)
XX
DE   Tomato/Escherichia coli chimeric TP-CPL DNA.
XX
KW   UDP-glucosyltransferase; pHBA; p-hydroxybenzoic acid ester glucoside;
KW   liquid crystal polymer; LCP; methylparaben; preservative; food;
KW   cosmetic industry; ds; gene; tomato; TP-CPL; chimeric;
KW   chorismate pyruvate lyase; plant.
XX
OS   Chimeric.
OS   Lycopersicon esculentum.
OS   Escherichia coli.
XX
```

```
FH Key      Location/Qualifiers
FT CDS      1..684
FT /tag= a
FT /product= "Tomato/Escherichia coli chimeric TP-CPL
FT protein"
XX
XX WO2003066836-A2.
XX
XX 14-AUG-2003.
XX
XX 06-FEB-2003; 2003WO-US005863.
XX
XX 07-FEB-2002; 2002US-0355511P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Meyer K, Van Dyk DE, Viitanen PV;
XX
XX WPI: 2003-767259/72.
XX
XX P-PSDB; ADC25989.
XX
XX New nucleic acid encoding UDP-glucosyltransferase, useful for preparing
XX cells that produce p-hydroxybenzoic acid glucose ester, also the new
XX enzymes.
XX
XX Example 9; SEQ ID NO 41; 161pp; English.
XX
XX The invention relates to a novel isolated nucleic acid that encodes a UDP
XX -glucosyltransferase. The method of the invention may be used to
XX transform microorganisms or green plant cells so that these produce a
XX higher level of high-volume chemicals or materials, such as pHBA (p-
XX hydroxybenzoic acid) ester glucoside which is an intermediate for pHBA, a
XX monomer for liquid crystal polymers (LCs) and starting material for
XX methylparaben. Methylparaben is a preservative commonly used in the food
XX and cosmetic industries. The encoded enzymes may be used for in vitro
XX production of these compounds and for identifying similar enzymes by
XX sequence comparison. The current sequence is that of the
XX tomato/Escherichia coli chimeric TP-CPL (chorismate pyruvate lyase) DNA
XX of the invention.
XX
XX Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;
XX
XX Query Match      100.0%; Score 684; DB 10; Length 684;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-219;
XX Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGCTTCTCTGCTGCTATTTCTTTCAGCAGCTGTTGCCACACGCAAGCAATGTTACACAAGCT 60
DB      1 ATGGCTTCTCTGCTGCTATTTCTTTCAGCAGCTGTTGCCACACGCAAGCAATGTTACACAAGCT 60

QY      61 AGCATGGTTGCACCTTTCTACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120
DB      61 AGCATGGTTGCACCTTTCTACTGCTCTCAAAATCTTCAGCCACTTTCCCTGTTACAAAGAAG 120

QY      121 CAAACCTTGACATCACTTCCATTTCTAGCAATGTTGGAGAGTTAGCTGCATGCAAGTG 180
DB      121 CAAACCTTGACATCACTTCCATTTCTAGCAATGTTGGAGAGTTAGCTGCATGCAAGTG 180

QY      181 TGGCATATGTCAACCCCGGTTAAACCAACTGCGTGCCTGCTATTGTAAGAAGATC 240
DB      181 TGGCATATGTCAACCCCGGTTAAACCAACTGCGTGCCTGCTATTGTAAGAAGATC 240

QY      241 CTGCGCTCGATCCGCAACTGCTGACTGCTGTTCTCGAGGATTTCCATGACAAACGT 300
DB      241 CTGCGCTCGATCCGCAACTGCTGACTGCTGTTCTCGAGGATTTCCATGACAAACGT 300

QY      301 TTTGAACACAGGAGAAACCGTGAACGATGATCCGCGAAGGTTTTCGAGCAG 360
DB      301 TTTGAACACAGGAGAAACCGTGAACGATGATCCGCGAAGGTTTTCGAGCAG 360

QY      361 AATGAAATCCCGAAGAACTGCGCGTGTGCGCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
DB      361 AATGAAATCCCGAAGAACTGCGCGTGTGCGCGAAAGAGTCTCGTTACTGGTTACGTGAA 420
```

QY 421 ATTTTGTATGTGCCGATGGTGAACCGTGGCTTGCCTGCTGCTACCGTGTTCCTGTCTCA 480
DB |||||||
QY 421 ATTTTGTATGTGCCGATGGTGAACCGTGGCTTGCCTGCTGCTACCGTGTTCCTGTCTCA 480
DB |||||||
QY 481 ACGTTAAGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACCGCTTAGGACGCTAT 540
DB |||||||
QY 481 ACGTTAAGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACCGCTTAGGACGCTAT 540
DB |||||||
QY 541 CTGTTACATCATCATGACATTAAACCGGACCTTATTGAGATAGGCGGTGATGCCGGGCTG 600
DB |||||||
QY 541 CTGTTACATCATCATGACATTAAACCGGACCTTATTGAGATAGGCGGTGATGCCGGGCTG 600
DB |||||||
QY 601 TGGGGGCGACGTTCCCGCTGCGATTAAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
DB |||||||
QY 601 TGGGGGCGACGTTCCCGCTGCGATTAAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
DB |||||||
QY 661 TTACCGGCGTCAACCGTTGTACTAA 684
DB |||||||

RESULT 4
ADV91622
ID ADV91622 standard; DNA; 684 BP.
XX
AC ADV91622;
XX
DT 10-MAR-2005 (first entry)
XX
DE Tomato rubisco small subunit precursor TP-E. coli CPL chimeric DNA.
XX
KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
KW HCHU; gene expression; hydroquinone; antiseptic; transgenic plant;
KW genetically engineered microorganism; antioxidant; antimicrobial agent;
KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
KW neoplasm; PHBA; para-hydroxybenzoic acid; p-hydroxybenzoic acid;
KW rubisco small subunit precursor; transit peptide;
KW chloroplast transit peptide; TP; chloroplast; gene; db.
XX
OS Lycopersicon esculentum.
OS Escherichia coli.
OS Chimeric.

Key Location/Qualifiers
CDS 1..684
FT /*tag= a
FT /product= "Tomato rubisco small subunit precursor TP-E.
FT coli CPL fusion protein"

US2004261147-A1.
XX
XX 23-DEC-2004.
XX
XX 16-JUN-2003; 2003US-00462162.
XX
XX 16-JUN-2003; 2003US-00462162.
XX
XX (MEYE/) MEYER K.
PA (VIIT/) VIITANEN P V.
PA (FLIN/) FLINT D.
XX
PI Meyer K, Viitanen PV, Flint D;
XX
XX WPI; 2005-057232/06.
DR P-PSDB; ADV91623.
XX

Producing hydroquinone glucoside in a green plant comprises growing a
PT green plant having nucleic acid fragments and chorismate pyruvate lyase
PT expression cassette.
XX
PS Example 4; SEQ ID NO 33; 70pp; English.

XX The present invention relates to methods and materials to produce
CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)
CC in genetically modified green plants and microorganisms. The method
CC relies upon transgenic plants or genetically modified microorganisms that
CC produce increased levels of the initial substrate para-hydroxybenzoic
CC acid (p-hydroxybenzoic acid; PHBA) in a biosynthetic pathway to produce
CC arbutin. Plants and microbes can be genetically engineered to produce
CC high levels of PHBA either by functional expression of the bacterial
CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-
CC hydroxycinnamoyl-CoA hydratase/lyase (HCHU) or a combination of both.
CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-
CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The
CC present sequence is tomato rubisco small subunit precursor (rbcs2)
CC transit peptide (chloroplast transit peptide; TP)-Escherichia coli CPL
CC chimeric DNA. This sequence is present in the E. coli expression vector
CC construct pET24a-TP-CPL.
XX
SQ Sequence 684 BP; 160 A; 171 C; 179 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 684; DB 14; Length 684;
Best Local Similarity 100.0%; Pred. No. 1.4e-219; Indels 0; Gaps 0;
Matches 684; Conservative 0; Mismatches 0;

QY 1 ATGGCTTCCTCTGTCACTTTCTTCAGCAGCTGTTCACACGACGCAATGTTACAAAGCT 60
DB |||||||
QY 1 ATGGCTTCCTCTGTCACTTTCTTCAGCAGCTGTTCACACGACGCAATGTTACAAAGCT 60
DB |||||||
QY 61 AGCATGGTTGACCTTTTCACTGGTCTCAATCTTCAGCAGCTTTCCCTGTTACAAAGAAG 120
DB |||||||
QY 61 AGCATGGTTGACCTTTTCACTGGTCTCAATCTTCAGCAGCTTTCCCTGTTACAAAGAAG 120
DB |||||||
QY 121 CAAACCTTGACATCAGCTTCCATTTAGTAACTGGTGAAGAGTTAGTGCATGACGAGTG 180
DB |||||||
QY 121 CAAACCTTGACATCAGCTTCCATTTAGTAACTGGTGAAGAGTTAGTGCATGACGAGTG 180
DB |||||||
QY 181 TGGCATATGTACACCCCGCTTAAACGCACTGGCTGCGCTATTGTTAAAGAGATC 240
DB |||||||
QY 181 TGGCATATGTACACCCCGCTTAAACGCACTGGCTGCGCTATTGTTAAAGAGATC 240
DB |||||||
QY 241 CTGCTGCGTGGATCGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB |||||||
QY 241 CTGCTGCGTGGATCGCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB |||||||
QY 301 TTGAAACAGCAGGAGGAAACCGTAAAGCGTACGATGATCCGGAAGGGTTTCTCGAGCAG 360
DB |||||||
QY 301 TTGAAACAGCAGGAGGAAACCGTAAAGCGTACGATGATCCGGAAGGGTTTCTCGAGCAG 360
DB |||||||
QY 361 AATGAAATCCCGAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB |||||||
QY 361 AATGAAATCCCGAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB |||||||
QY 421 ATTTTGTATGTGCCGATGGTGAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB |||||||
QY 421 ATTTTGTATGTGCCGATGGTGAACCGTGGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB |||||||
QY 481 ACGTTAAGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACCGCTTAGGACGCTAT 540
DB |||||||
QY 481 ACGTTAAGCGGCGGAGCTGGCGTTACAAAATTGGGTAAACCGCTTAGGACGCTAT 540
DB |||||||
QY 541 CTGTTACATCATCATGACATTAAACCGGACCTTATTGAGATAGGCGGTGATGCCGGGCTG 600
DB |||||||
QY 541 CTGTTACATCATCATGACATTAAACCGGACCTTATTGAGATAGGCGGTGATGCCGGGCTG 600
DB |||||||
QY 601 TGGGGGCGACGTTCCCGCTGCGATTAAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
DB |||||||
QY 601 TGGGGGCGACGTTCCCGCTGCGATTAAACCGGTAACCGCTGTTGCTTAACAGAACTGTTT 660
DB |||||||
QY 661 TTACCGGCGTCAACCGTTGTACTAA 684
DB |||||||
QY 661 TTACCGGCGTCAACCGTTGTACTAA 684
DB |||||||

RESULT 5
ADV91636
ID ADV91636 standard; DNA; 3452 BP.
XX AC ADV91636;
XX AC ADV91636;
XX DT 10-MAR-2005 (first entry)
XX DE C. parapsilosis pHB 1-H- E. coli CPL- A. thaliana UGT72B1 chimeric DNA.
XX KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
XX KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
XX KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;
XX KW genetically engineered microorganism; antioxidant; antimicrobial agent;
XX KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
XX KW neoplasm; pHBA 1-hydroxylase; pHBA 1-H; pHBA; para-hydroxybenzoic acid;
XX KW p-hydroxybenzoic acid; UDP-glucosyltransferase; UGT72B1; gene; ds.
XX OS Arabidopsis thaliana.
XX OS Candida parapsilosis.
XX OS Escherichia coli.
XX OS Chimeric.
XX FH Key Location/Qualifiers
XX FT CDS 4..501
XX FT FT /*tag= a
XX FT /product= "Escherichia coli CPL protein"
XX FT CDS 541..1980
XX FT FT /*tag= b
XX FT /product= "Candida parapsilosis pHBA 1-H protein"
XX FT /transl_except= (pos: 559..561, aa:Trp)
XX FT CDS 2004..3446
XX FT FT /*tag= c
XX FT /product= "Arabidopsis thaliana UGT72B1 protein"
XX US2004261147-A1.
XX PN
XX XX
XX PD 23-DEC-2004.
XX PF 16-JUN-2003; 2003US-00462162.
XX PR 16-JUN-2003; 2003US-00462162.
XX XX (MEYER/) MEYER K.
XX PA (VIIT/) VIITANEN P V.
XX PA (FLIN/) FLINT D.
XX PI Meyer K, Viitanen PV, Flint D;
XX XX
XX DR WPI; 2005-057232/06.
XX DR P-PSDB; ADV91605, ADV91619, ADV91631.
XX PT Producing hydroquinone glucoside in a green plant comprises growing a
XX PT green plant having nucleic acid fragments and chorismate pyruvate lyase
XX PT expression cassette.
XX PS Example 7; SEQ ID NO 47; 70pp; English.
XX CC The present invention relates to methods and materials to produce
XX CC hydroquinone glucoside (arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside)
XX CC in genetically modified green plants and microorganisms. The method
XX CC relies upon transgenic plants or genetically modified microorganisms that
XX CC produce increased levels of the initial substrate para-hydroxybenzoic
XX CC acid (p-hydroxybenzoic acid; pHBA) in a biosynthetic pathway to produce
XX CC arbutin. Plants and microbes can be genetically engineered to produce
XX CC high levels of pHBA either by functional expression of the bacterial
XX CC protein chorismate pyruvate lyase (CPL) or by the expression of 4-
XX CC hydroxycinnamoyl-CoA hydratase/lyase (HCHL) or a combination of both.
XX CC Arbutin is useful as an antioxidant, antimicrobial agent, anti-
XX CC inflammatory agent and inhibitor of carcinogenesis (melanoma). The
XX CC present sequence is Candida parapsilosis pHBA 1-hydroxylase (pHB 1-H) -
XX CC Escherichia coli chorismate pyruvate lyase (CPL) and Arabidopsis thaliana
XX CC UDP-glucosyltransferase (UGT72B1) chimeric DNA.

XX SQ Sequence 3452 BP; 932 A; 730 C; 856 G; 934 T; 0 U; 0 Other;
XX Query Match 73.2%; Score 501; DB 14; Length 3452;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-157;
XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 184 CATATGTCACCCCGGTTAACGCAACTGCGTGCCTGCGCTATTGTAAGAGATCCCT 243
DB 1 CATATGTCACCCCGGTTAACGCAACTGCGTGCCTGCGCTATTGTAAGAGATCCCT 60
QY 244 GCCTCGATCCGCAACTGCTCGACTGCTGCTGCGAGATTCCCATGACAAACGTTTT 303
DB 61 GCCTCGATCCGCAACTGCTCGACTGCTGCTGCGAGATTCCCATGACAAACGTTTT 120
QY 304 GAACAGCAGGAGAAAACGGTAAAGCTGACGATGATCCGGAAGGGTTTGTGAGCAGAAT 363
DB 121 GAACAGCAGGAGAAAACGGTAAAGCTGACGATGATCCGGAAGGGTTTGTGAGCAGAAT 180
QY 364 GAAATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTGTTACTGTTACGTGAAATT 423
DB 181 GAAATCCCGAAGAACTGCGCTGCTGCCGAAGAGTCTGTTACTGTTACGTGAAATT 240
QY 424 TTGTTATGTCGCGATGGTGAACCGTGTGCGGTGCTACCGTCCGTTCCTGTGTCAACG 483
DB 241 TTGTTATGTCGCGATGGTGAACCGTGTGCGGTGCTACCGTCCGTTCCTGTGTCAACG 300
QY 484 TTAAGCGCGCGAGCTGGCTTACAAAATTGGGTAAACGCGGTAGACGCTATCTG 543
DB 301 TTAAGCGCGCGAGCTGGCTTACAAAATTGGGTAAACGCGGTAGACGCTATCTG 360
QY 544 TTCACATCATGACATTAAACCGGACCTTTATTGAGATAGGCCGCGTGTGGTGG 603
DB 361 TTCACATCATGACATTAAACCGGACCTTTATTGAGATAGGCCGCGTGTGGTGG 420
QY 604 GGGCGACGTTCCCGCTCGGATTAAACCGGTAAACCGCTGTGCTTAACAGAACTGTTTTA 663
DB 421 GGGCGACGTTCCCGCTCGGATTAAACCGGTAAACCGCTGTGCTTAACAGAACTGTTTTA 480
QY 664 CCGGCGTCACCGTTGTACTAA 684
DB 481 CCGGCGTCACCGTTGTACTAA 501
RESULT 6
ADV91643
ID ADV91643 standard; DNA; 1971 BP.
XX AC ADV91643;
XX DT 10-MAR-2005 (first entry)
XX DE C. parapsilosis pHB 1-H- E. coli CPL chimeric DNA.
XX KW Hydroquinone glucoside; arbutin; 4-hydroxyphenyl-beta-D-glucopyranoside;
XX KW chorismate pyruvate lyase; CPL; 4-hydroxycinnamoyl-CoA hydratase/lyase;
XX KW HCHL; gene expression; hydroquinone; antiseptic; transgenic plant;
XX KW genetically engineered microorganism; antioxidant; antimicrobial agent;
XX KW anti-inflammatory agent; carcinogenic inhibitor; melanoma; cytostatic;
XX KW neoplasm; pHBA 1-hydroxylase; pHBA; para-hydroxybenzoic acid;
XX KW p-hydroxybenzoic acid; gene; ds.
XX OS Candida parapsilosis.
XX OS Escherichia coli.
XX OS Chimeric.
XX FH Key Location/Qualifiers
XX FT CDS 4..501
XX FT FT /*tag= a
XX FT /product= "Escherichia coli CPL protein"
XX FT /transl_except= (pos: 46..48, aa:Cys)
XX FT CDS 522..1561
XX FT FT /*tag= b


```
CC pET24a(+).
XX Sequence 498 BP; 114 A; 122 C; 141 G; 121 T; 0 U; 0 Other;
SQ

Query Match 72.8%; Score 498; DB 14; Length 498;
Best Local Similarity 100.0%; Pred. No. 7.1e-157;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCAACCCCGGTTAAGCAACTGCGTGCCTGCGCTATTGTAAGAGATCCCTGCC 246
Db 1 ATGTCAACCCCGGTTAAGCAACTGCGTGCCTGCGCTATTGTAAGAGATCCCTGCC 60

QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTCTGGAGATTCATGACAAACGTTTGA 306
Db 61 CTGGATCCGCAACTGCTGCACTGGCTGTTCTGGAGATTCATGACAAACGTTTGA 120

QY 307 CAGCAGGGAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAGAAATGA 366
Db 121 CAGCAGGGAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAGAAATGA 180

QY 367 ATCCCGAAGAACTGCGCTGCTGCGGAAAGAGTCTGTTACTGTTACGTGAAATTTG 426
Db 181 ATCCCGAAGAACTGCGCTGCTGCGGAAAGAGTCTGTTACTGTTACGTGAAATTTG 240

QY 427 TTATGTCCGATGGTGAACCGTGGCTTGCCTGCGGTACCGTCTGTTCTGTGTCAACGTTA 486
Db 241 TTATGTCCGATGGTGAACCGTGGCTTGCCTGCGGTACCGTCTGTTCTGTGTCAACGTTA 300

QY 487 AGCGGCGCGAGCTGGCTTTACAAAATTGGGTAAACGCGTTAGGACGCTATCTGTTT 546
Db 301 AGCGGCGCGAGCTGGCTTTACAAAATTGGGTAAACGCGTTAGGACGCTATCTGTTT 360

QY 547 ACATCATGACATTAACCCGGGACTTATTGAGATAGGCGCGTGAATCGCGGCTGTGGGG 606
Db 361 ACATCATGACATTAACCCGGGACTTATTGAGATAGGCGCGTGAATCGCGGCTGTGGGG 420

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTACAGAACTGTTTACCG 666
Db 421 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTACAGAACTGTTTACCG 480

QY 667 GCGTCACCGTTGACTAA 684
Db 481 GCGTCACCGTTGACTAA 498

RESULT 8
AAD06956
ID AAD06956 standard; DNA; 854 BP.
XX
AC AAD06956;
XX
XX 06-AUG-2001 (first entry)
DT
DE ubiC gene encoding chorismate pyruvate lyase.
XX
XX P-hydroxybenzoic acid; pHb; aromatic pathway; carbon source; ubiC;
KW chorismate pyruvate lyase; ppsA; phosphoenolpyruvate synthase;
KW PEP synthase; tktA; transketolase; aroC; chorismate synthase;
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSP synthase; aroA;
KW shikimate kinase II; aroB; 3-dehydroquininate synthase; DHQ synthase; aroG;
KW 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; ds.
XX
OS Escherichia coli.
XX
XX US6210937-B1.
XX
XX 03-APR-2001.
XX
XX 22-APR-1998; 98US-00064693.
XX
XX 22-APR-1997; 97US-0044094P.
XX
XX (BECH-) BECHTEL BWXT IDAHO LLC.
PA
```

```
XX Ward TE, Watkins CS, Bulmer DK, Johnson BF, Amaratunga M;
XX WPI; 2001-280857/29.
XX
XX Producing aromatic compounds, especially commercially acceptable levels
XX of p-hydroxybenzoic acid, comprises developing genetically engineered
XX bacteria that carry selected genes of the common aromatic pathway.
XX
XX Claim 5; Col 31-32; 25pp; English.
XX
XX The present invention relates to a method of producing p-hydroxybenzoic
XX acid (pHB) which comprises transforming microorganisms with plasmids
XX carrying selected genes of the common aromatic pathway and by conversion
XX of carbon sources. The selected genes of the common aromatic pathway of
XX the invention includes ubiC gene encoding chorismate pyruvate lyase, aroG
XX structural gene encoding 3-deoxy-D-arabino-heptulosonate-7-phosphate
XX (DHAP) isoenzyme synthase (phe), tktA gene encoding transketolase, ppsA
XX gene encoding phosphoenolpyruvate (PEP) synthase, aroC gene encoding
XX chorismate synthase, aroL gene encoding shikimate kinase II, aroA gene
XX encoding 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase and aroB
XX gene encoding 3-dehydroquininate (DHQ) synthase. The method is used for
XX converting chorismate to a selected aromatic compound. The method is
XX particularly used for the production of commercially acceptable levels of
XX aromatic compounds, particularly p-hydroxybenzoic acid. The present
XX sequence is ubiC gene which encodes chorismate pyruvate lyase that
XX catalyses the production of chorismate from pHb
XX
SQ Sequence 854 BP; 202 A; 200 C; 222 G; 230 T; 0 U; 0 Other;

Query Match 72.8%; Score 498; DB 4; Length 854;
Best Local Similarity 100.0%; Pred. No. 9.4e-157;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTCAACCCCGGTTAAGCAACTGCGTGCCTGCGCTATTGTAAGAGATCCCTGCC 246
Db 321 ATGTCAACCCCGGTTAAGCAACTGCGTGCCTGCGCTATTGTAAGAGATCCCTGCC 380

QY 247 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 306
Db 381 CTGGATCCGCAACTGCTGCACTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 440

QY 307 CAGCAGGGAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAGAAATGA 366
Db 441 CAGCAGGGAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAGAAATGA 500

QY 367 ATCCCGAAGAACTGCGCTGCTGCGGAAAGAGTCTGTTACTGTTTACGTGAAATTTG 426
Db 501 ATCCCGAAGAACTGCGCTGCTGCGGAAAGAGTCTGTTACTGTTTACGTGAAATTTG 560

QY 427 TTATGTCCGATGGTGAACCGTGGCTTGCCTGCGGTACCGTCTGTTCTGTGTCAACGTTA 486
Db 561 TTATGTCCGATGGTGAACCGTGGCTTGCCTGCGGTACCGTCTGTTCTGTGTCAACGTTA 620

QY 487 AGCGGCGCGAGCTGGCTTTACAAAATTGGGTAAACGCGTTAGGACGCTATCTGTTT 546
Db 621 AGCGGCGCGAGCTGGCTTTACAAAATTGGGTAAACGCGTTAGGACGCTATCTGTTT 680

QY 547 ACATCATGACATTAACCCGGGACTTATTGAGATAGGCGCGTGAATCGCGGCTGTGGGG 606
Db 681 ACATCATGACATTAACCCGGGACTTATTGAGATAGGCGCGTGAATCGCGGCTGTGGGG 740

QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTACAGAACTGTTTACCG 666
Db 741 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTACAGAACTGTTTACCG 800

QY 667 GCGTCACCGTTGACTAA 684
Db 801 GCGTCACCGTTGACTAA 818

RESULT 9
AAD06957
```


XX	
PR	30-JUN-1994; 94DE-04423022.
XX	
XX	(HEID/) HEIDE L.
PA	
XX	Heide L, Siebert M, Severin K;
XX	
PI	
XX	WPI; 1995-186908/25.
DR	P-PSDB; AAR74742.
DR	

XX Claim 3: Page 5: 7pp; German.

Transgenic plants containing a bacterial gene which encodes an enzyme able to catalyse prodn. of an antiviral, bactericidal, fungicidal or insecticidal factor are new. A preferred gene is *ubiC* from *E. coli* (AAQ92409) which codes for chorismate-pyruvate lyase (AAR74742). The lyase catalyses conversion of chorismate to p-hydroxybenzoic acid and transgenic tobacco plants which express the active enzyme are resistant to tobacco mosaic virus

SQ Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;
Query Match 72.4%; Score 495; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 7.2e-156;

Qy 187 ATGTCACCCCGGTTAAGCGCAACTCGTCGCGCTGGCTATTGTAAAGAGATCCCTGCC 2

1 ATGTCACACCCCGGTTAACGCAACTGCGTGGCTATTGTTAAAGAGATCCTGCC

Accession	Sequence	Position
24/09	CTGGATCCGCAACATGCTCGACATGCGTGTGCTGGAGGATTCATGACACAAACGTTTTGAA	1
Db	61 CTGGATCCGCAACATGCTCGACATGCGTGTGCTGGAGGATTCATGACACAAACGTTTTGAA	1
09	307 CAGCAGCGGAAAAACGGTAMAGCTTCACGATGATCCCGGAAGGTTTGTTCGAGCAGAAATGAA	1

Qy

307 CAGCAGGGAAAAACGGTTAAGCGTCACGATGATCCGGAAGGGTTTGTCTGAGCAGAATACTAA
D6

121 CAGCAGGGAAAAACGGTTAAGCGTCACGATGATCCGGAAGGGTTTGTCTGAGCAGAATACTAA

Qy
Db

367 ATCCCCGAAGAACTCGCGTCTGCCGAAAGAGTCGTACTGGTTACGTGAATTTTG
181 ATCCCCGAAGAACTCGCGTCTGCCGAAAGAGTCGTACTGGTTACGTGAATTTTG

[illegible][illegible]

Db 301 AGCGGCGGAGCTGGCGTTACAAAATTGGGTAAAAACGCCGTTAGGACGCTTATCTGTTC

Qy 547 ACATCATCGACATTAACCCCGGAGCTTTATTGAGATAGCCCGTGATCGCGGGGTGTGGGGG

361 ACATCATCGACATTAAACCGGACTTTATTGAGATAGCCGCTGATGCCGGGCTGTGGGGG
607 CGAGTTCCCGCCTGGCATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTATTACCG

421 CGAGTTCCCGCCCTGCGATTAAACCGTGTTCCTAACAGAACTGTTTACCG
667 GCGTCACGGTTGTAC 681

QY 66 / GCGTCACGGTTGTAC 681
|||||
481 GCGTCACGGTTGTAC 495

RESULT 12
ABA91837
ID ABA91837 standard; DNA; 495 BP.

10 ABA91837 standard; DNA; 495 bp.
XX
XX
AC ABA91837;
XX

DT	15-MAY-2002 (first entry)	DT	61	CTGGATCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA	120
XX	Escherichia coli chorismate pyruvate lyase gene coding region.	Db			
DE		QY	307	CAGCAGGAAAAACGGTTAAGCGTACGATGATCCGGAAGGTTTGTTCGACGAGATGAA	366
KW	Chorismate pyruvate lyase; CPL; enzyme; p-hydroxybenzoic acid;	Db	121	CAGCAGGAAAAACGGTTAAGCGTACGATGATCCGGAAGGTTTGTTCGACGAGATGAA	180
KW	transgenic plant; ubiC; gene; ds.	QY	367	ATCCCCGAAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTTGAAATTTTG	426
XX	Escherichia coli.	Db	181	ATCCCCGAAAGAACTGCCGCTGCTGCCGAAAGAGTCTCGTTACTGTTGAAATTTTG	240
XX		QY	427	TTATGTGCCGATGTGAACCGTTCGCCGCTGCTACCGTCTGTTCTGTGTCAACGTTA	486
FH	Key	Db	241	TTATGTGCCGATGTGAACCGTTCGCCGCTGCTACCGTCTGTTCTGTGTCAACGTTA	300
FT	CDS	QY	487	AGCGGGCGGAGCTGGCGTTACAAAATTGGGTAACCGCTTAGGACGCTATCTGTTTC	546
FT	Location/Qualifiers	Db	301	AGCGGGCGGAGCTGGCGTTACAAAATTGGGTAACCGCTTAGGACGCTATCTGTTTC	360
FT	1..495	QY	547	ACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCCGCTGATGCCGGCTGTGGGG	606
FT	/*tag= a	Db	361	ACATCATCGACATTAACCCGGGACTTTATTGAGATAGGCCGCTGATGCCGGCTGTGGGG	420
FT	/partial	QY	607	CGAGTTCCCGCCCTGCGATTAAAGCGGTAAACCGCTGTGCTTAACAGAACTGTTTTACCG	666
FT	/product= "Chorismate pyruvate lyase"	Db	421	CGAGTTCCCGCCCTGCGATTAAAGCGGTAAACCGCTGTGCTTAACAGAACTGTTTTACCG	480
FT	/note= "the CDS does not include a stop codon"	QY	667	GCCTCACCGTTGTAC	681
XX	WO200194607-A2.	Db	481	GCCTCACCGTTGTAC	495
XX	13-DEC-2001.				
XX	22-MAY-2001; 2001WO-US016661.				
XX	02-JUN-2000; 2000US-0209854P.				
XX	(DUPO) DU PONT DE NEMOURS & CO E I.				
XX	Meyer K, Van Dyk DE, Viitanen PV;				
XX	WPI; 2002-226795/28.				
DR	P-PSDB; AAM50958.				
XX					
PT	Producing para-hydroxy benzoic acid in green plant, comprises expression				
PT	of unique expression cassette containing gene encoding chorismate				
PT	pyruvate lyase operably linked to specific chloroplast targeting				
PT	sequence.				
XX					
PS	Claim 4; Page 55; 60pp; English.				
XX					
CC	The present sequence is that of the coding region of the Escherichia coli				
CC	strain W3110 ubiC gene in expression construct pET24a. The gene encodes				
CC	chorismate pyruvate lyase (CPL, see AAM50968). It was obtained by PCR				
CC	amplification of strain W3110 genomic DNA using primers (see ABA91839-40)				
CC	based on the published E. coli ubiC gene. The CPL open reading frame is				
CC	used in an expression cassette designed for the high-level production of				
CC	p-hydroxybenzoic acid (pHBA) in green plants. The expression cassette				
CC	comprises the CPL coding sequence operably linked to a promoter capable				
CC	of driving protein expression in higher plants. The cassette also has a				
CC	sequence encoding a chloroplast transit peptide, its natural cleavage				
CC	site, and a small portion of a transit peptide donor protein fused to the				
CC	N-terminus of CPL. The chloroplast targeting sequence targets the foreign				
CC	protein to the chloroplast and aids in its uptake into the organelle. The				
CC	cleavage site is unique to the transit peptide, and cleavage of the				
CC	chimeric protein encoded by the cassette at this site releases a novel				
CC	polypeptide that has full enzyme activity, comprising the mature CPL				
CC	enzyme and a small portion of transit peptide donor (see AAM50961). A				
CC	plant comprising the CPL expression cassette is claimed, and may be				
CC	soybean, rapeseed, sunflower, cotton, corn, tobacco, alfalfa, wheat,				
CC	barley, oats, sorghum, rice, Arabidopsis, sugarbeet, sugarcane, canola,				
CC	millet, bean, pea, rye, flax or a forage grass. pHBA is a monomeric				
CC	component of liquid crystal polymers which have application in the				
CC	automotive, electrical and other industries				
XX					
SQ	Sequence 495 BP; 112 A; 122 C; 141 G; 120 T; 0 U; 0 Other;				
	Query Match				
	Best Local Similarity 72.4%; Score 495; DB 6; Length 495;				
	Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	187 ATGTCACACCCCGCGTTAACGCAACTGCGTGCCTATTGTGTAAGAGATCCCTGCC	246			
Db	1 ATGTCACACCCCGCGTTAACGCAACTGCGTGCCTATTGTGTAAGAGATCCCTGCC	60			
QY	247 CTGGATCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA	306			

Producing aromatic acid conjugates, involves contacting glycosylated

301 AGCGGCCGGAGCTGGCGTTACAAAAATTGGGTAAACGCCGTTAGGACGCTATCTGTTC 360

101

KW cosme

liquid crystal polymer; LCP; mesomorphic liquid crystalline polyimide; cosmesic industry; dg. gene.

liquid crystal polymer; LC

irradiation; ds; gene; gene; ubiC; CPL; chorismate synthase; mefenoxam; ds; gene; ubiC; CPL; chorismate synthase

22

QY 547 ACATCATCGACATTAAACCGGACCTTTATTTGAGATAGCCGCTGATCCCGGCTGTGGGG 606
Dd |||||
Dd 361 ACATCATCGACATTAAACCGGACCTTTATTTGAGATAGCCGCTGATCCCGGCTGTGGGG 420
QY 607 CGACGTTCCCGCTCGGATTAAAGCGTAAACCGCTGTTCTAACAGAACTGTTTTTACCG 666
Dd |||||
Dd 421 CGACGTTCCCGCTCGGATTAAAGCGTAAACCGCTGTTCTAACAGAACTGTTTTTACCG 480
QY 667 GCGTCACCGTTGTAC 681
Dd |||||
Dd 481 GCGTCACCGTTGTAC 495

RESULT 15

AAS90335
ID AAS90335 standard; cDNA; 1207 BP.

XX AC
AAS90335;

XX DT
13-FEB-2002 (first entry)

XX DE
DNA encoding novel human diagnostic protein #26139.

XX KW
Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS
Homo sapiens.

XX PN
WO200175067-A2.

XX PD
11-OCT-2001.

XX PF
30-MAR-2001; 2001WO-US008631.

XX PR
31-MAR-2000; 2000US-00540217.

XX PR
23-AUG-2000; 2000US-00649167.

XX PA
(HYSE-) HYSEQ INC.

XX PI
Drmanac RT, Liu C, Tang YT;

XX DR
WPI; 2001-639362/73.

XX DR
P-PSDB; ABG26148.

XX PT
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.

XX PS
Claim 1; SEQ ID NO 26139; 103bp; English.

XX CC
The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (II) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
coding sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1207 BP; 316 A; 265 C; 288 G; 338 T; 0 U; 0 Other;

Query Match 69.1%; Score 472.8; DB 5; Length 1207;
Best Local Similarity 99.2%; Pred. No. 3.6e-148;
Matches 496; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 187 ATGTCACACCCCGGTTAAACGCAACTGCGTGGCTGGCTATTGTAAAGAGATCCCTGCC 246

Dd |||||

QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGTGGAGAGATTCATGACAAAAAGTTTGA 306

Dd |||||

QY 307 CAGCAGGAAAAACGGTTAAGCGTGACGATGATCCGCGAAGGTTTGTGAGCAGAAATG 366

Dd |||||

QY 367 ATCCCGGAAGAACTGCGCTGCTCCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTG 426

Dd |||||

QY 427 T-TATGTGCCGATGGTGAACCGTGGCTTGCC-GGTGCTACCGTCTGTTCTGTGCAACGT 484

Dd |||||

QY 485 TAAGCGGCCGAGCTGGCGTTACAAAATTTGGGTAAACCGCGTTAGGACGCTATCTGT 544

Dd |||||

QY 545 TCACATCATCGACATTAAACCGGACCTTTATTGAGATAGGCGGTGATCCCGGCTGTGGG 604

Dd |||||

QY 605 GCGACGTTCCCGCTCGGATTAAAGCGGTAACCGCTGTTGCTAACAGAACTGTTTTTAC 664

Dd |||||

QY 665 CGGCGTCACCGTTGTACTAA 684

Dd |||||

QY 1132 CGGCGTCACCGTTGTACTAA 1151

Search completed: July 11, 2006, 22:41:55
Job time : 590 secs

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 11, 2006, 19:07:37 ; Search time 4394 Seconds
(without alignments)
9954.492 Million cell updates/sec
Title: US-10-718-311-7
Perfect score: 684
Sequence: 1 atggcttctctgtcatttc.....cggcgtaaccgttgtactaa 684
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 6366136 seqs, 31973710525 residues
Total number of hits satisfying chosen parameters: 12732272
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_env.*
2: gb_pat.*
3: gb_ph.*
4: gb_pl.*
5: gb_pr.*
6: gb_ro.*
7: gb_sts.*
8: gb_sy.*
9: gb_un.*
10: gb_vi.*
11: gb_ov.*
12: gb_htg.*
13: gb_in.*
14: gb_om.*
15: gb_ba.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	684	100.0	684 2 AR428736 Sequence
2	684	100.0	684 2 AR455412 Sequence
3	684	100.0	684 2 AX329368 Sequence
4	684	100.0	684 2 AX382258 Sequence
5	498	72.8	854 2 AR144788 Sequence
6	498	72.8	1498 15 ECUUBIA M96268 Escherichia
7	498	72.8	2000 2 E11274 ubiC and ub
8	498	72.8	2000 15 ECUUBICA M93136 Escherichia
9	498	72.8	2000 15 ECUUBIPLS M93413 Escherichia
10	498	72.8	2034 15 ECUUBIAC X57434 E.coli ubiC
11	498	72.8	2348 15 ECUUBI X66619 E.coli gene
12	498	72.8	6641 2 AR144789 Sequence
13	498	72.8	110000 15 U00096_42 Continuation (43 o
14	498	72.8	110000 15 AP009048_42 U00006 E. coli chr
15	498	72.8	176195 15 ECUUBI89 DQ087228 Escherich
16	496.4	72.6	1396 15 DQ087228
17	495	72.4	495 2 A47932 Sequence 1
18	495	72.4	495 2 AR428733 Sequence

19	495	72.4	495	2	AR455409	AR455409 Sequence
20	495	72.4	495	2	AX329364	AX329364 Sequence
21	495	72.4	495	2	AX382254	Continuation (42 o
22	494.8	72.3	110000	15	CP000036_41	Continuation (52 o
23	490	71.6	110000	15	AE005174_51	Continuation (51 o
24	488.4	71.4	110000	15	BA000007_50	Continuation (52 o
25	488.4	71.4	110000	15	BA000007_51	Continuation (43 o
26	488.4	71.4	110000	15	CP000034_42	Continuation (45 o
27	488.4	71.4	110000	15	CP000038_44	Continuation (48 o
28	485.2	70.9	110000	15	AE014075_47	E07849 gDNA encodi
29	481.2	70.4	1918	2	E07849	Continuation (44 o
30	478.8	70.0	110000	15	AE005674_43	Continuation (35 o
31	478.8	70.0	110000	15	AE014073_34	AE008898 Salmonell
32	320.8	46.9	23880	15	AE008898	Continuation (43 o
33	319.2	46.7	110000	15	AE014613_42	Continuation (44 o
34	319.2	46.7	110000	15	AE017220_43	Continuation (42 o
35	319.2	46.7	110000	15	CP000026_41	Continuation (43 o
36	319.2	46.7	110000	15	CP000026_42	AL627282 Salmonell
37	319.2	46.7	234050	15	AL627282	AR384505 Sequence
38	282	41.2	645	2	AR384505	M13543 Tomato (L.e
39	184	26.9	599	4	TOMRBCSB	M15236 Tomato RuBP
40	184	26.9	742	4	TOMRBCSE	BT013023 Lycopersi
41	184	26.9	1097	4	BT013023	X05983 Tomato rbcS
42	179	26.2	2776	4	LERBCS2	X69759 S. tuberosum
43	146.8	21.5	3323	4	STRBCS1	AY220079 Nicotiana
44	145.6	21.3	806	4	AY220079	X05985 Tomato rbcS
45	143.6	21.0	1054	4	LERBCS3B	

ALIGNMENTS

RESULT 1
AR428736
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN
Query Match
Best Local Similarity
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCTCTGTGTCATTTCTTTCAGCAGCTGTTTCCACACGCGCAGCAATGTTACACAAGCT 60
DB 1 ATGGCTTCTCTGTGTCATTTCTTTCAGCAGCTGTTTCCACACGCGCAGCAATGTTACACAAGCT 60
QY 61 AGCATGGTTGACCTTCTACTGGTCTCAAAATCTTACGCCACTTTCCCTGTTTCAAGAAG 120
DB 61 AGCATGGTTGACCTTCTACTGGTCTCAAAATCTTACGCCACTTTCCCTGTTTCAAGAAG 120
QY 121 CAAAACCTTGACATCACTTCCATGCTAGCAATGTTGGAAGAGTTCGCTGCAGGCTG 180
DB 121 CAAAACCTTGACATCACTTCCATGCTAGCAATGTTGGAAGAGTTCGCTGCAGGCTG 180
QY 181 TGGCATATGTCACACCCCGCTTAACTGCGTGCCTGCGCTATTTGTAAGAAGATC 240

```
Db      181  TGGCATATGTCTACACCCCGGTTAAACGCAACTGCGTGGCGCTATTGTTAAAGAGATC 240
Qy      241  CTTGCCCTGGATCCGCAACTGCTCGACTGGCTGGTGTGGAGGATTCCATGACAAAACGT 300
Db      241  CTTGCCCTGGATCCGCAACTGCTCGACTGGCTGGTGTGGAGGATTCCATGACAAAACGT 300
Qy      301  TTTTGAACAGCAGGAGGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360
Db      301  TTTTGAACAGCAGGAGGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360
Qy      361  AATGAAATCCCGGAAAGAACTGCCGTGCTGCCGGAAGAGTCTCGTTACTGGTTACGTGAA 420
Db      361  AATGAAATCCCGGAAAGAACTGCCGTGCTGCCGGAAGAGTCTCGTTACTGGTTACGTGAA 420
Qy      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGC CGGTGCTGACCGTCTCTCTGTGTCA 480
Db      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGC CGGTGCTGACCGTCTCTCTGTGTCA 480
Qy      481  ACGTTTAAAGCGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCCGTTAGGACGCTAT 540
Db      481  ACGTTTAAAGCGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCCGTTAGGACGCTAT 540
Qy      541  CTGTTTCATCATCATGACATTAACCCGGGACTTTTATGAGATAGGCGCGTGATGCCGGGCTG 600
Db      541  CTGTTTCATCATCATGACATTAACCCGGGACTTTTATGAGATAGGCGCGTGATGCCGGGCTG 600
Qy      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy      661  TTACCGGCGTCAACCGTTGTACTAA 684
Db      661  TTACCGGCGTCAACCGTTGTACTAA 684
```

```
RESULT 2
LOCUS      AR455412                      684 bp      DNA      linear      PAT 20-FEB-2004
DEFINITION Sequence 7 from patent US 6683231.
ACCESSION  AR455412
VERSION     AR455412.1  GI:42689960
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 684)
AUTHORS     Meyer, K., Viitanen, P.V. and Van Dyk, D.E.
TITLE       High level production of p-hydroxybenzoic acid in green plants
JOURNAL     Patent: US 6683231-A 7 27-JAN-2004;
            E. I. du Pont de Nemours and Company; Wilmington, DE
FEATURES    Location/Qualifiers
             source
               1..684
               /organism="unknown"
               /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 684; DB 2; Length 684;
Best Local Similarity 100.0%; Pred. No. 6.8e-202; Mismatches 0; Indels 0; Gaps 0;
Matches 684; Conservative 0;

Qy      1  ATGGCTTCTCTGTCTGATTTCTTTACAGAGCTGTGTCACACGAGCAATGTTACACAAGCT 60
Db      1  ATGGCTTCTCTGTCTGATTTCTTTACAGAGCTGTGTCACACGAGCAATGTTACACAAGCT 60
Qy      61  AGCATGGTGGACCTTCACTGGTCTCAAAATCTTCCAGCCACTTTCCCTGTTTACAAAGAG 120
Db      61  AGCATGGTGGACCTTCACTGGTCTCAAAATCTTCCAGCCACTTTCCCTGTTTACAAAGAG 120
Qy      121  CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG 180
Db      121  CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG 180
Qy      181  TGGCATATGTCTACACCCCGGTTAAACGCAACTGCGTGGCGCTATTGTTAAAGAGATC 240
```

```
Db      181  TGGCATATGTCTACACCCCGGTTAAACGCAACTGCGTGGCGCTATTGTTAAAGAGATC 240
Qy      241  CTTGCCCTGGATCCGCAACTGCTCGACTGGCTGGTGTGGAGGATTCCATGACAAAACGT 300
Db      241  CTTGCCCTGGATCCGCAACTGCTCGACTGGCTGGTGTGGAGGATTCCATGACAAAACGT 300
Qy      301  TTTTGAACAGCAGGAGGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360
Db      301  TTTTGAACAGCAGGAGGAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTCCAGCAG 360
Qy      361  AATGAAATCCCGGAAAGAACTGCCGTGCTGCCGGAAGAGTCTCGTTACTGGTTACGTGAA 420
Db      361  AATGAAATCCCGGAAAGAACTGCCGTGCTGCCGGAAGAGTCTCGTTACTGGTTACGTGAA 420
Qy      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGC CGGTGCTGACCGTCTCTCTGTGTCA 480
Db      421  ATTTTGTATTATGTCGCGATGGTGAACCGTGGCTTGC CGGTGCTGACCGTCTCTCTGTGTCA 480
Qy      481  ACGTTTAAAGCGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCCGTTAGGACGCTAT 540
Db      481  ACGTTTAAAGCGGCGCGAGCTGGCGTTACAAAATTTGGGTAAACCGCCGTTAGGACGCTAT 540
Qy      541  CTGTTTCATCATCATGACATTAACCCGGGACTTTTATGAGATAGGCGCGTGATGCCGGGCTG 600
Db      541  CTGTTTCATCATCATGACATTAACCCGGGACTTTTATGAGATAGGCGCGTGATGCCGGGCTG 600
Qy      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db      601  TGGGGGCGACGTTCCCGCTCGGATTAACCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Qy      661  TTACCGGCGTCAACCGTTGTACTAA 684
Db      661  TTACCGGCGTCAACCGTTGTACTAA 684

RESULT 3
LOCUS      AX329368                      684 bp      DNA      linear      PAT 08-JAN-2002
DEFINITION Sequence 7 from Patent WO0194607.
ACCESSION  AX329368
VERSION     AX329368.1  GI:18102367
KEYWORDS    .
SOURCE      synthetic construct
ORGANISM    synthetic construct
REFERENCE   1
AUTHORS     Meyer, K., van Dyk, D.E. and Viitanen, P.V.
TITLE       High level production of p-hydroxybenzoic acid in green plants
JOURNAL     Patent: WO 0194607-A 7 13-DEC-2001;
            E.I. Dupont De Nemours (US)
FEATURES    Location/Qualifiers
             source
               1..684
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="synthetic CPL"

ORIGIN
Query Match      100.0%; Score 684; DB 2; Length 684;
Best Local Similarity 100.0%; Pred. No. 6.8e-202; Mismatches 0; Indels 0; Gaps 0;
Matches 684; Conservative 0;

Qy      1  ATGGCTTCTCTGTCTGATTTCTTTACAGAGCTGTGTCACACGAGCAATGTTACACAAGCT 60
Db      1  ATGGCTTCTCTGTCTGATTTCTTTACAGAGCTGTGTCACACGAGCAATGTTACACAAGCT 60
Qy      61  AGCATGGTGGACCTTCACTGGTCTCAAAATCTTCCAGCCACTTTCCCTGTTTACAAAGAG 120
Db      61  AGCATGGTGGACCTTCACTGGTCTCAAAATCTTCCAGCCACTTTCCCTGTTTACAAAGAG 120
Qy      121  CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG 180
Db      121  CAAAACCTTGACATCACTTCCATTGCTAGCAATGGTGAAGAGTTAGCTGCATGCAAGTG 180
```

Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGTCGATCAGAGTG 180
QY 181 TGGCATATGTCAACCCCGGTTAACGAACTGCGTGCCTATTTGTAAGAGATC 240
Db 181 TGGCATATGTCAACCCCGGTTAACGAACTGCGTGCCTATTTGTAAGAGATC 240
QY 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGAGAGATTCATGACAAACGT 300
Db 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGAGAGATTCATGACAAACGT 300
QY 301 TTTGAACAGCAGGAGAAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAG 360
Db 301 TTTGAACAGCAGGAGAAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAG 360
QY 361 AATGAATCCCGAAGAACTGCGCTGCTGCGAAGAGATTCGTTACTGGTTACGTGAA 420
Db 361 AATGAATCCCGAAGAACTGCGCTGCTGCGAAGAGATTCGTTACTGGTTACGTGAA 420
QY 421 ATTTTGTATGTGCGCATGTTGAAACCGTGTGCGGTCGTACCGTCTGCTGTGTC 480
Db 421 ATTTTGTATGTGCGCATGTTGAAACCGTGTGCGGTCGTACCGTCTGCTGTGTC 480
QY 481 ACGTTAAGCGGCGGAGCTGGGTTACAAAATTCGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACGTTAAGCGGCGGAGCTGGGTTACAAAATTCGGTAAACCGCGTTAGGACGCTAT 540
QY 541 CTGTTACATCATCGACATTAACCGGAGCTTTATTGAGATAGCGCGTATGCGGGCTG 600
Db 541 CTGTTACATCATCGACATTAACCGGAGCTTTATTGAGATAGCGCGTATGCGGGCTG 600
QY 601 TGGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
QY 661 TTACCGGCGTCACCGTTGTACTAA 684
Db 661 TTACCGGCGTCACCGTTGTACTAA 684

RESULT 4
LOCUS AX382258 684 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 16 from Patent WO0204653.
ACCESSION AX382258
VERSION AX382258.1 GI:19577037
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Flint, D., Meyer, K. and Viitanen, P.V.
TITLE Sinapolyglucose:malate sinapolytransferase form malate conjugates
JOURNAL from benzoic acid glucosides
E.I. DUPONT DE NEMOURS AND COMPANY (US)
FEATURES
source
1. .684
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="open reading frame of the chloroplast-targeted CPL fusion protein"

ORIGIN
Query Match 100.0%; Score 684; DB 2; Length 684;
Best Local Similarity 100.0%; Pred. No. 6.8e-202;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTGTCACACGCGAGCAATGTCACAGCT 60
Db 1 ATGGCTTCTCTGTCATTTCTTTCAGCAGCTGTGTCACACGCGAGCAATGTCACAGCT 60
QY 61 AGCATGTTGCACCTTTCACTGCTCAAAATCTTCAGCCACTTTTCCCTGTGTACAAAGAG 120

Db 61 AGCATGTTGCACCTTTCACTGCTCAAAATCTTCAGCCACTTTTCCCTGTGTACAAAGAG 120
QY 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGTCGATCAGAGTG 180
Db 121 CAAACCTTGACATCACTTCCATTGCTAGCAATGGTGGAGAGTTAGTCGATCAGAGTG 180
QY 181 TGGCATATGTCAACCCCGGTTAACGAACTGCGTGCCTATTTGTAAGAGATC 240
Db 181 TGGCATATGTCAACCCCGGTTAACGAACTGCGTGCCTATTTGTAAGAGATC 240
QY 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGAGAGATTCATGACAAACGT 300
Db 241 CTGCGCTGGATCCGCAACTGCTCGACTGGCTGCTGAGAGATTCATGACAAACGT 300
QY 301 TTTGAACAGCAGGAGAAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAG 360
Db 301 TTTGAACAGCAGGAGAAACCGGTAAAGCGTGAAGATCCGCGAAGGGTTTGTGAGCAG 360
QY 361 AATGAATCCCGAAGAACTGCGCTGCTGCGAAGAGATTCGTTACTGGTTACGTGAA 420
Db 361 AATGAATCCCGAAGAACTGCGCTGCTGCGAAGAGATTCGTTACTGGTTACGTGAA 420
QY 421 ATTTTGTATGTGCGCATGTTGAAACCGTGTGCGGTCGTACCGTCTGCTGTGTC 480
Db 421 ATTTTGTATGTGCGCATGTTGAAACCGTGTGCGGTCGTACCGTCTGCTGTGTC 480
QY 481 ACGTTAAGCGGCGGAGCTGGGTTACAAAATTCGGTAAACCGCGTTAGGACGCTAT 540
Db 481 ACGTTAAGCGGCGGAGCTGGGTTACAAAATTCGGTAAACCGCGTTAGGACGCTAT 540
QY 541 CTGTTACATCATCGACATTAACCGGAGCTTTATTGAGATAGCGCGTATGCGGGCTG 600
Db 541 CTGTTACATCATCGACATTAACCGGAGCTTTATTGAGATAGCGCGTATGCGGGCTG 600
QY 601 TGGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
Db 601 TGGGGGCGAGCTTCCCGCTGCGATTAAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTT 660
QY 661 TTACCGGCGTCACCGTTGTACTAA 684
Db 661 TTACCGGCGTCACCGTTGTACTAA 684

RESULT 5
LOCUS AR144788 854 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 24 from patent US 6210937.
ACCESSION AR144788
VERSION AR144788.1 GI:15106655
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 854)
AUTHORS Ward, T.E., Watkins, C.S., Bulmer, D.K., Johnson, B.F. and Amaralunga, M.
TITLE Development of genetically engineered bacteria for production of selected aromatic compounds
JOURNAL Patent: US 6210937-A 24 03-APR-2001;
FEATURES Location/Qualifiers
source
1. .854
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 72.8%; Score 498; DB 2; Length 854;
Best Local Similarity 100.0%; Pred. No. 9.5e-144;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 ATGTACACCCCGGTTAACGCAACTGCGTGCCTGCGCTATTTGTAAGAGATCCCTGCC 246
Db 321 ATGTACACCCCGGTTAACGCAACTGCGTGCCTGCGCTATTTGTAAGAGATCCCTGCC 380

QY 247 CTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAAAGCTTTTGA 306
Db |||||
QY 381 CTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGGATTCATGACAAAAAGCTTTTGA 440
Db |||||
QY 307 CAGCAGGAAAAACGGTAAGCTGACGATGATCCCGAAGGGTTTGTCCAGCAGAAATGA 366
Db |||||
QY 441 CAGCAGGAAAAACGGTAAGCTGACGATGATCCCGAAGGGTTTGTCCAGCAGAAATGA 500
QY 367 ATCCCGAGAACTCCGCTGCTCCGGAAGAGTCTGTTACTGTTACGTTAGTGAATTTTG 426
Db |||||
QY 501 ATCCCGAAGAACTCCGCTGCTCCGGAAGAGTCTGTTACTGTTACGTTAGTGAATTTTG 560
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTGCCGCTGCTACCGTCTGTTCTCTGTGTCAACGTTA 486
Db |||||
QY 561 TTATGTGCCGATGGTGAACCGTGGCTTGCCGCTGCTACCGTCTGTTCTGTGTCAACGTTA 620
QY 487 AGCGGGCCGAGCTGGCTTACAAAAATTGGGTAAAAACGCCGTTAGGACGCTATCTGTTTC 546
Db |||||
QY 621 AGCGGGCCGAGCTGGCTTACAAAAATTGGGTAAAAACGCCGTTAGGACGCTATCTGTTTC 680
QY 547 ACATCATCGACATTAACCGGAGCTTTATTGAGATAGGCGCTGATGCCGGCTGTGGGG 606
Db |||||
QY 681 ACATCATCGACATTAACCGGAGCTTTATTGAGATAGGCGCTGATGCCGGCTGTGGGG 740
QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTTACCG 666
Db |||||
QY 741 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTTACCG 800
QY 667 GCGTCACCGTTGTACTAA 684
Db |||||
QY 801 GCGTCACCGTTGTACTAA 818

RESULT 6
ECOUBIA 1498 bp DNA linear BCT 28-WAR-1994
LOCUS Escherichia coli 4-hydroxybenzoate octaprenyl transferase (ubiA)
DEFINITION gene, 5' end and chorismate lyase (ubiC) gene, complete cds.
ACCESSION M96268
VERSION M96268.1 GI:347886
KEYWORDS 4-hydroxybenzoate octaprenyl transferase; aerobic respiratory deficiency; chorismate lyase; membrane-bound protein; prenyltransferase; transferase; ubiA gene; ubiC gene; ubiquinone biosynthesis.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 1498)
AUTHORS Wu, G., Williams, H. D., Gibson, F. and Poole, R. K.
TITLE Mutants of Escherichia coli affected in respiration: the cloning and nucleotide sequence of ubiA, encoding the membrane-bound p-hydroxybenzoate:octaprenyltransferase
JOURNAL J. Gen. Microbiol. 139 (pt 8), 1795-1805 (1993)
PUBMED 840922
FEATURES Location/Qualifiers
source 1..1498
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/strains="K-12"
/db_xref="taxon:562"
/map="91.8 min"
/tissue_lib="cosmid"
41..524
/note="ORF; the sequence before nucleotide number 110 might not be genuine. Three Clai sites, which is the enzyme used for cloning, were identified in this region, so it could be a cloning artifact"
/codon_start=3
/transl_table=11
/protein_id="AAA17026.1"
/db_xref="GI:347887"

/translation="RFSEVNPDTQOMIDFINRENTNKWLIDKNNNFFIQLIDQALRST
DDMIKANVHLYKEMIRSDVSPIFETEDNLRTFNELTRNDNIFLFSVDDGPV
MYVSSQRLHDLMIAPTKDTNWNSTIYKSRHEMLPVNLTOETLFPSSKSHGKXALPIFT
ASWRAHRIMNKGV"
735..738
RBS
gene 747..1244
/gene="ubiC"
CDS 747..1244
/gene="ubiC"
/codon_start=1
/transl_table=11
/product="chorismate lyase"
/protein_id="AAA17027.1"
/db_xref="GI:347888"
/translation="MSHPALTOURLALRYCKEIPALDPOLLWLLEDSMTKRFEOQK
TVSTMIREGFVEQNEIPEELPLPKESRYWLREILLACDGEPLAGTIVVPVSTLSG
PELAQLKLGKTPGLRYLFTSILTRDFIEIGRDAGLWRRRLSLSGKPLLLTELFLP
ASPLY"
1257..1498
gene /gene="ubiA"
CDS 1257..1498
/gene="ubiA"
/note="The last part of this ORF is to be found in entry K00127.; ORF"
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA17028.1"
/db_xref="GI:347889"
/translation="MEWSLTQNKLLAFHRLMRTDKPIGALLLLWPTLLWALWVATPGVP
QLWTLAVFVAGVWLMRAAGCVVNDYADRFEDGHVKRT"

ORIGIN
Query Match 72.8%; Score 498; DB 15; Length 1498;
Best Local Similarity 100.0%; Pred. No. 1.1e-143; Mismatches 0; Gaps 0;
Matches 498; Conservative 0; Indels 0;
QY 187 ATGTCACACCCCGCGGTTAAACGCAACTGCGTGGCTATTTGTAAGAGATCCCTGCC 246
Db |||||
QY 247 CTGATCCGCAACTGCTGCGTGGCTGCTGAGAGATCCATGACAAAAAGCTTTTGA 306
Db |||||
QY 307 CAGCAGGAAAAACGGTTAAGCGTCACGATGATCCCGAAGGGTTTGTCCAGCAGAAATGA 366
Db |||||
QY 367 ATCCCGAAGAACTGCCGCTGCTCCGAAAGAGTCTCGTTACTGGTTACGTGAATTTTG 426
Db |||||
QY 427 TTATGTCCGATGTGAAACCGTGGCTGCTCCCGGTGCTACCGTCTGTTCTGTGTCAACGTTA 486
Db |||||
QY 487 AGCGGGCCGAGCTGGCGTTACAAAAATTGGTAAAAACGCCGTTAGGACGCTATCTGTTTC 546
Db |||||
QY 547 ACATCATCGACATTAACCCGGGACTTTATTGAGATAGCCGCTGATGCCGGCTGTGGGG 606
Db |||||
QY 607 CGAGCTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAACAGAACTGTTTTACCG 666
Db |||||
QY 667 GCGTCACCGTTGTACTAA 684
Db |||||
QY 1227 GCGTCACCGTTGTACTAA 1244

```

RESULT 7
E11274
LOCUS      ubiC and ubiA gene.                2000 bp    DNA    linear    PAT 29-SEP-1997
DEFINITION
ACCESSION  E11274
VERSION    E11274.1 GI:22024916
KEYWORDS   JP 1996107789-A/1.
SOURCE     Escherichia coli
ORGANISM   Bacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE  1 (bases 1 to 2000)
AUTHORS    Matsuda,H., Kawamuki,M. and Nakagawa,T.
TITLE      PRODUCTION OF UBIQUINONE-10
JOURNAL    Patent: JP 1996107789-A 1 30-APR-1996;
            ALPHA- SHOKUHIN KK
COMMENT     OS Escherichia coli
            PN JP 1996107789-A/1
            PD 30-APR-1996
            PF 13-OCT-1994 JP 1994273071
            PI MATSUDA HIDEYUKI, KAWAMUKI MAKOTO, NAKAGAWA TSUYOSHI PC
            C12N15/09,A61K31/12,A61K31/12,C07H21/04,C12N1/21,C12P7/66, PC
            (C12N1/21,
            PC C12R1:01), (C12P7/66,C12R1:01);
            CC strandedness: Double;
            CC topology: Linear;
            FH Key Location/Qualifiers
            FH source 1. .2000
            FT /organism='Escherichia coli'
            FT CDS 380..877
            FT /product='UbiC protein'
            FT CDS 890..1762
            FT /product='UbiA protein'.
            FT Location/Qualifiers
            FT 1. .2000
            /organism='Escherichia coli'
            /mol_type='genomic DNA'
            /db_xref='taxon:562'

ORIGIN
Query Match 72.8%; Score 498; DB 2; Length 2000;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGTACACCCCGCTTAACGCAACTCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 246
DB 187 ATGTACACCCCGCTTAACGCAACTCGTGGCTGCGCTATTGTAAGAGATCCCTGCC 246
QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTCTCGAGGATTCATGACAAACGTTTGA 306
DB 247 CTGGATCCGCAACTGCTCGACTGGCTGTTCTCGAGGATTCATGACAAACGTTTGA 306
QY 440 CTGGATCCGCAACTGCTCGACTGGCTGTTCTCGAGGATTCATGACAAACGTTTGA 499
DB 440 CTGGATCCGCAACTGCTCGACTGGCTGTTCTCGAGGATTCATGACAAACGTTTGA 499
QY 307 CAGCAGGAAAAACGTAAGCTGACATGATCCGGAAGGTTTCTCGAGCAGATGA 366
DB 307 CAGCAGGAAAAACGTAAGCTGACATGATCCGGAAGGTTTCTCGAGCAGATGA 366
QY 500 CAGCAGGAAAAACGTAAGCTGACATGATCCGGAAGGTTTCTCGAGCAGATGA 559
DB 500 CAGCAGGAAAAACGTAAGCTGACATGATCCGGAAGGTTTCTCGAGCAGATGA 559
QY 367 ATCCCGAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGGTTTACGTGAAATTTG 426
DB 367 ATCCCGAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGGTTTACGTGAAATTTG 426
QY 560 ATCCCGAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGGTTTACGTGAAATTTG 619
DB 560 ATCCCGAGAACTGCGCTGCTGCCGAAAGAGTCTGTTACTGGTTTACGTGAAATTTG 619
QY 427 TTATGTGCCGATGGTGAACCGTGGCTTTCGGGTCGTACCGTTCGTGTCGCAACGTTA 486
DB 427 TTATGTGCCGATGGTGAACCGTGGCTTTCGGGTCGTACCGTTCGTGTCGCAACGTTA 486
QY 620 TTATGTGCCGATGGTGAACCGTGGCTTTCGGGTCGTACCGTTCGTGTCGCAACGTTA 679
DB 620 TTATGTGCCGATGGTGAACCGTGGCTTTCGGGTCGTACCGTTCGTGTCGCAACGTTA 679
QY 487 AGCGGCGGAGCTGGGTTACAAAATTTGGGTAAACCGCGTTAGACCTATCTGTTTC 546
DB 487 AGCGGCGGAGCTGGGTTACAAAATTTGGGTAAACCGCGTTAGACCTATCTGTTTC 546
QY 680 AGCGGCGGAGCTGGGTTACAAAATTTGGGTAAACCGCGTTAGACCTATCTGTTTC 739
DB 680 AGCGGCGGAGCTGGGTTACAAAATTTGGGTAAACCGCGTTAGACCTATCTGTTTC 739
QY 547 ACATCATCGACATTAAACCGGGACTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 606
DB 547 ACATCATCGACATTAAACCGGGACTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 606
QY 740 ACATCATCGACATTAAACCGGGACTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 799
DB 740 ACATCATCGACATTAAACCGGGACTTTATTGAGATAGGCGGTGATGCCGGCTGTGGGG 799

```

```

QY 607 CGACGTTCCCGCTCGATTAAACGCGTAAACCGCTGCTGCTAACAGAACTGTTTTACCG 666
DB 800 CGACGTTCCCGCTCGATTAAACGCGTAAACCGCTGCTGCTAACAGAACTGTTTTACCG 859
QY 667 GCGTCACCGTTGTAATAA 684
DB 860 GCGTCACCGTTGTAATAA 877

RESULT 8
ECOUBICA      2000 bp    DNA    linear    BCT 26-APR-1993
LOCUS      Escherichia coli chorismate lyase (ubiC), 4-hydroxybenzoate
DEFINITION  octaprenyl transferase (ubiA) genes, complete cds, and
            sn-glycerol-3-phosphate acyltransferase (plsB) genes, 3' end.
ACCESSION  M93136
VERSION    M93136.1 GI:148099
KEYWORDS   4-hydroxybenzoate octaprenyl transferase; chorismate lyase; plsB
            gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC
            gene.
SOURCE     Escherichia coli W3110
ORGANISM   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Escherichia.
REFERENCE  1 (bases 1 to 2000)
AUTHORS    Lightner,V.A., Bell,R.M. and Modrich,P.
TITLE      The DNA sequences encoding plsB and dgk loci of Escherichia coli
JOURNAL    J. Biol. Chem. 258 (18), 10856-10861 (1983)
COMMENT     6309817
            2 (bases 1 to 2000)
            Nichols,B.P. and Green,J.M.
            Cloning and Sequencing of Escherichia coli ubiC and purification of
            chorismate lyase
            Unpublished (1992)
            Original source text: Escherichia coli (sub_strain W3110, strain
            K-12) (library: Kohare lambda-lf8) DNA.
FEATURES    Location/Qualifiers
            1. .2000
            /organism='Escherichia coli W3110'
            /mol_type='genomic DNA'
            /strain='K-12'
            /sub_strain='W3110'
            /db_xref='taxon:316407'
            /map_xref='91.8 min'
            /tissue_lib='Kohare lambda-lf8'
            380..877
            /gene='ubiC'
            380..877
            /gene='ubiC'
            /function='enzymatic chorismate --> p-hydroxybenzoate +
            pyruvate'
            /experiment='experimental evidence, no additional details
            recorded'
            /citation=[2]
            /codon_start=1
            /transl_table=11
            /product='chorismate lyase'
            /protein_id='AAA24711.1'
            /db_xref='GI:148100'
            /translation='MSHPALTOLRALRYCKEIPALDPOLLMDLLEDSTMTRFEQOQK
            TVSVNMRGPFVEQNEIPEELLLPKESRYWLEILLCADGEPMLAGRTVVPVSTLSG
            PELALQKUGKPLGLYFTSTLTDFIIGRDAGLWGRRLRLSGKPLLLTFLFPL
            ASPLA'
            890..1762
            /gene='ubiA'
            890..1762
            /gene='ubiA'
            /function='enzymatic - 3-octaprenyl-4-hydroxybenzoate
            synthesis'
            /experiment='experimental evidence, no additional details
            recorded'
            /citation=[2]

```

```
gene complete cds, chorisamate lyase (ubiC) gene complete cds,
sn-glycerol-3-phosphate acyltransferase (plsB) gene, 3' end.
M93413.1 GI:148106
M93413.1 GI:148106
4-hydroxybenzoate octaprenyl transferase; chorisamate lyase; plsB
gene; sn-glycerol-3-phosphate acyltransferase; ubiA gene; ubiC
gene.
Escherichia coli W3110
Escherichia coli W3110
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 2000)
Lightner, V.A., Bell, R.M. and Modrich, P.
The DNA sequences encoding plsB and dgk loci of Escherichia coli
J. Biol. Chem. 258 (18), 10856-10861 (1983)
6309817
REFERENCE 2 (bases 1 to 2000)
Nichols, B.P. and Green, J.M.
Cloning and sequencing of Escherichia coli ubiC and purification of
chorisamate lyase
J. Bacteriol. 174 (16), 5309-5316 (1992)
1644758
COMMENT
Original source text: Escherichia coli (sub_strain W3110, strain
K-12) (library: Kohara lambda) DNA.
Location/Qualifiers
1. 2000
/organism="Escherichia coli W3110"
/mol_type="genomic DNA"
/strain="K-12"
/sub_strain="W3110"
/db_xref="taxon:316407"
/map="91.8 min"
/tissue_lib="Kohara lambda"
380..877
/gene="ubiC"
380..877
/gene="ubiC"
/function="enzymatic chorisamate --> p-hydroxybenzoate +
pyruvate"
/experiment="experimental evidence, no additional details
recorded"
/feature="N-terminal amino acid sequence of chorisamate lyase
: NH2-SHPALTQALRYCEIPAL-"
/citation=[2]
/codon_start=1
/transl_table=11
/product="chorisamate lyase"
/db_xref="GI:148107"
/transl_table="MSPHALTQALRYCKEIPALDQDLDLWLLDMSMTKRFQOQK
TVSTMIREGFEQNEIPEELPLPKESRYLWREILLCADGEPWLAGTVVPVSTLSG
PELALOKLGTPLGRYLFSTSLTRDIEIGRDAGLWGRSRLSLSGKPLLLTELFLP
ASPLY"
890..1762
/gene="ubiA"
890..1762
/gene="ubiA"
/function="enzymatic - 3-octaprenyl-4-hydroxybenzoate
synthesis"
/experiment="experimental evidence, no additional details
recorded"
/citation=[2]
/number=2
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate-octaprenyl transferase"
/protein_id="AAA24717.1"
/db_xref="GI:148108"
/transl_table="MEWSLTQNKLLAFHRLMRTDKPICALLLWPTLWLVATPGVP
QLWILAVFVAGVWLMRAAGCVANDYADKFDGHVKRTANPLPSPGAVTEKEARALFVV
LVLSIFLVLTNTMTLLSIAALAWVYPMKRYTHLPQVVLGAAGWSIPMAFAA
VSESVPFLSCHLMFLANILWAVDYOYAMVDKDKIKSTAILFGDYDKLILGIL
QIGVLALMIISELNGLGWYWSILVAGALFYQOKLIANRERACFAFMNNIVG
LVFLGLAWSYWHF"
complement (1917..2000)
/gene="plsB"
complement (1917..2000)
/gene="plsB"
/EC_number="2.3.1.15"
/function="lysophosphatidic acid synthesis"
/experiment="experimental evidence, no additional details
recorded"
/citation=[1]
/codon_start=1
/transl_table=11
/product="sn-glycerol-3-phosphate acyltransferase"
/protein_id="AAA24713.1"
/db_xref="GI:148102"
/transl_table="MKVYQLLAELITSDVRLTIESATQEG"
```

ORIGIN

```
Query Match 72.8%; Score 498; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 187 ATGTACACACCCGCTTAACGCAACTCGTCGCTGCGCTATTGTAAGAGATCCCTGCC 246
DB 380 ATGTACACACCCGCTTAACGCAACTCGTCGCTGCGCTATTGTAAGAGATCCCTGCC 439
QY 247 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 306
DB 440 CTGGATCCGCAACTGCTCGACTGGCTGTTGCTGGAGGATTCATGACAAAACGTTTGA 499
QY 307 CAGCAGGGAACCGTAAGCGTGAAGATGATCCGGAAGGTTTGTGCGAGCAGATGAA 366
DB 500 CAGCAGGGAACCGTAAGCGTGAAGATGATCCGGAAGGTTTGTGCGAGCAGATGAA 559
QY 367 ATCCCGGAAGACTCGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAAATTTTG 426
DB 560 ATCCCGGAAGACTCGCGCTGCTGCCGAAGAGTCTCGTTACTGTTACGTGAAATTTTG 619
QY 427 TTATGTCCGATGGTGAACCGTGGCTTCCGCGTCTGATCCGTCGTTCTGTGTCACGTTA 486
DB 620 TTATGTCCGATGGTGAACCGTGGCTTCCGCGTCTGATCCGTCGTTCTGTGTCACGTTA 679
QY 487 AGCGGGCGGAGCTCGCGTTACAAAATTTGGTAAACCGCGTTAGGACGCTATCTGTC 546
DB 680 AGCGGGCGGAGCTCGCGTTACAAAATTTGGTAAACCGCGTTAGGACGCTATCTGTC 739
QY 547 ACATCATCGACATTAACCCGGGACTTTATTGATAGCGCGTATGATCCGCGGCTGTGGGG 606
DB 740 ACATCATCGACATTAACCCGGGACTTTATTGATAGCGCGTATGATCCGCGGCTGTGGGG 799
QY 607 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTACC 666
DB 800 CGACGTTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTACC 859
QY 667 GCGTCACCGTTGTACTAA 684
DB 860 GCGTCACCGTTGTACTAA 877
```

RESULT 9

```
ECOUBIPLS Escherichia coli 4-hydroxybenzoate octaprenyl transferase (ubiA)
LOCUS 2000 bp DNA linear BCT 26-APR-1993
DEFINITION
```

```
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
COMMENT
FEATURES
source
```

gene

CDS

gene

CDS

gene LVLFLGLAMSYWHP
complement (1917. .2000)
/gene="pleB"
CDS complement (1917. .2000)
/gene="pleB"
/EC number="2.3.1.15"
/function="lysophosphatidic acid synthesis"
/experiment="experimental evidence, no additional details recorded"
/citation={1}
/codon_start=1
/transl_table=11
/product="sn-glycerol-3-phosphate acyltransferase"
/protein_id="AA24718.1"
/db_xref="GI:148109"
/translation="MKVYQLAELITSVRLTIESATQEGG"

ORIGIN

Query Match 72.8%; Score 498; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred. No. 1.2e-143; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 0;

QY 187 ATGTACACCCCGGTTAAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 246
DB 380 ATGTACACCCCGGTTAAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 439
QY 247 CTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 306
DB 440 CTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 499
QY 307 CAGCAGGGAACGTAAGCGTACGATGATCCGGAAGGTTTCTCGAGCAGATGAA 366
DB 500 CAGCAGGGAACGTAAGCGTACGATGATCCGGAAGGTTTCTCGAGCAGATGAA 559
QY 367 ATCCCCGAAGAACTGCGGCTGCTGCCGAAGAGTCTGTTACTGTTAGCTGAAATTTG 426
DB 560 ATCCCCGAAGAACTGCGGCTGCTGCCGAAGAGTCTGTTACTGTTAGCTGAAATTTG 619
QY 427 TTATGTCCGATGTTGAACCGTGGCTGTTGCGGTCGCTGCTGCTGCTGCTGCTGCTG 486
DB 620 TTATGTCCGATGTTGAACCGTGGCTGTTGCGGTCGCTGCTGCTGCTGCTGCTGCTG 679
QY 487 AGCGGCGGAGCTGCGGTTACAAAATTTGGTTAAACCGCTTAGCGCTATCGTTTC 546
DB 680 AGCGGCGGAGCTGCGGTTACAAAATTTGGTTAAACCGCTTAGCGCTATCGTTTC 739
QY 547 ACATCATGACATTACCGGAGCTTTATTGATAGGCGGCTGATCCGGGCTGTGGGG 606
DB 740 ACATCATGACATTACCGGAGCTTTATTGATAGGCGGCTGATCCGGGCTGTGGGG 799
QY 607 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTACCG 666
DB 800 CGACGTTCCCGCTGCGATTAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTACCG 859
QY 667 GCGTCACCGTTGACTAA 684
DB 860 GCGTCACCGTTGACTAA 877

RESULT 10
LOCUS ECUBIAC 2034 bp DNA linear BCT 05-OCT-1992
DEFINITION E.coli ubiC and ubiA genes for chorismate lyase and 4-hydroxybenzoate octaprenyltransferase.
ACCESSION X57434
VERSION X57434.1 GI:43233
KEYWORDS 4-hydroxybenzoate octaprenyl transferase; 4-hydroxybenzoate synthetase; chorismate lyase.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

REFERENCE 1

AUTHORS Nishimura, K., Nakahigashi, K. and Inokuchi, H.
TITLE Location of the ubiA gene on the physical map of Escherichia coli
J. Bacteriol. 174 (17), 5762 (1992)
PUBMED 1512213
REFERENCE 2 (bases 1 to 2034)
AUTHORS Nishimura, K.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1991) K. Nishimura, Faculty of Sciences, Kyoto University, Sakyo-ku, Kyoto 606, Japan
FEATURES
Location/Qualifiers
1. .2034
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="K12 W3110"
/db_xref="taxon:562"
/map="91.7 min"
/clone="634"
/clone_lib="Kohara library"
-35 signal 137. .142
-10 signal 159. .164
gene 291. .899
/gene="ubiC"
CDS 291. .899
/gene="ubiC"
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate synthetase"
/protein_id="CAA40681.1"
/db_xref="GI:43234"
/db_xref="GOA:P26602"
/db_xref="UniProtKB/Swiss-Prot:P26602"
translation="MRLRPFCCVLHLCFTSPVNTFLRYNAFTLNCBFGMSHPALT QURALYKBIIPALDPQLDLLEDLSMTKRFQGGKTVSTMIRGFVEQNSIPBEL PLPLKESRYLWREILLCADGEPLAGRTVVPVSTLSPPELALQKLGKTLPLGYLFTSS TLTROFIEIGRDAGLWGRSRRLSGKPLLLTELFLPASPLY"
912. .1784
/gene="ubiA"
912. .1784
/gene="ubiA"
/codon_start=1
/transl_table=11
/product="4-hydroxybenzoate octaprenyltransferase"
/protein_id="CAA40682.1"
/db_xref="GI:43235"
/db_xref="GOA:P26601"
/db_xref="UniProtKB/Swiss-Prot:P26601"
translation="MEWSLTQNKLLAFHRLMRDTPIGALLLLWPTLWALWVATPGVP OLWILAVFVAGVWLMRAAGCVNDVADRFKDHVKRTPANRPLPSGAVTEKEARALFV LVLJSLFLVLTNTMTILLSIAALAWVYPEWKRYTHLPQVVLGAAPGWSIPMAFAA VSESVPLSCHWLMFLANIWNAYDTQYAMVDRDDVKIGIKSTAILFGQYDKLIIGIL QIGVLAALWIIIGELNGWGYYSILVAGALFVYQOKLIANRERERACFKAFMNNVTVG LVLFLGLAMSYWHP"

ORIGIN

Query Match 72.8%; Score 498; DB 15; Length 2034;
Best Local Similarity 100.0%; Pred. No. 1.2e-143; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 0;

QY 187 ATGTACACCCCGGTTAAACGCAACTCGCTGCGCTGCTATTTGTAAGAGATCCCTGCC 246
DB 402 ATGTACACCCCGGTTAAACGCAACTCGCTGCGCTATTTGTAAGAGATCCCTGCC 461
QY 247 CTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 306
DB 462 CTGGATCCGCAACTGCTGACCTGGCTGTTGCTGGAGATTCATGACAAACGTTTGA 521
QY 307 CAGCAGGGAACGTAAGCGTACGATGATCCGGAAGGTTTCTCGAGCAGATGAA 366
DB 522 CAGCAGGGAACGTAAGCGTACGATGATCCGGAAGGTTTCTCGAGCAGATGAA 581
QY 367 ATCCCCGAAGAACTGCGGCTGCTGCCGAAGAGTCTGTTACTGTTAGCTGAAATTTG 426
DB 582 ATCCCCGAAGAACTGCGGCTGCTGCCGAAGAGTCTGTTACTGTTAGCTGAAATTTG 641

Db 50949 CGAGCTCCCGCTCGGATTAAGCGGTAAACCGCTGTTGCTTAACAGAACTGTTTTTACCG 51008
Qy 667 GCGTCACCGTTGTACTAA 684
Db 51009 GCGTCACCGTTGTACTAA 51026

RESULT 14
AP009048_42
WPCOMMENT

Sequence split into 47 fragments LOCUS AP009048 Accession AP009048

Fragment Name Begin End

AP009048_00 1 110000
AP009048_01 100001 210000
AP009048_02 200001 310000
AP009048_03 300001 410000
AP009048_04 400001 510000
AP009048_05 500001 610000
AP009048_06 600001 710000
AP009048_07 700001 810000
AP009048_08 800001 910000
AP009048_09 900001 1010000
AP009048_10 1000001 1110000
AP009048_11 1100001 1210000
AP009048_12 1200001 1310000
AP009048_13 1300001 1410000
AP009048_14 1400001 1510000
AP009048_15 1500001 1610000
AP009048_16 1600001 1710000
AP009048_17 1700001 1810000
AP009048_18 1800001 1910000
AP009048_19 1900001 2010000
AP009048_20 2000001 2110000
AP009048_21 2100001 2210000
AP009048_22 2200001 2310000
AP009048_23 2300001 2410000
AP009048_24 2400001 2510000
AP009048_25 2500001 2610000
AP009048_26 2600001 2710000
AP009048_27 2700001 2810000
AP009048_28 2800001 2910000
AP009048_29 2900001 3010000
AP009048_30 3000001 3110000
AP009048_31 3100001 3210000
AP009048_32 3200001 3310000
AP009048_33 3300001 3410000
AP009048_34 3400001 3510000
AP009048_35 3500001 3610000
AP009048_36 3600001 3710000
AP009048_37 3700001 3810000
AP009048_38 3800001 3910000
AP009048_39 3900001 4010000
AP009048_40 4000001 4110000
AP009048_41 4100001 4210000
AP009048_42 4200001 4310000
AP009048_43 4300001 4410000
AP009048_44 4400001 4510000
AP009048_45 4500001 4610000
AP009048_46 4600001 4646332
Continuation (43 of 47) of AP009048 from base 4200001 AP009048 Escherichia coli W3110

Query Match 72.8%; Score 498; DB 15; Length 110000;

Best Local Similarity 100.0%; Pred. No. 3e-143;

Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 187 ATGTACACCCCGCTTAACGCAACTGCGTGGCGCTGATGTAAGAGATCCCTGCC 246
Db 56096 ATGTACACCCCGCGCTTAACGCAACTGCGTGGCGCTATGTAAGAGATCCCTGCC 56155
Qy 247 CTGGATCCGCAACTGCTGCACCTGGCTGTTGCTGGAGGATTCATCACAAAACGTTTGA 306
Db 56156 CTGGATCCGCAACTGCTGCACCTGGCTGTTGCTGGAGGATTCATCACAAAACGTTTGA 56215

Qy 307 CAGCAGGAAAAACGGTAAGCGTGACGATGATCCGGAAGGGTTTGTTCGAGCAGATGAA 366
Db 56216 CAGCAGGAAAAACGGTAAGCGTGACGATGATCCGGAAGGGTTTGTTCGAGCAGATGAA 56275
Qy 367 ATCCCGGAAGAACTGCGCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAATTTTG 426
Db 56276 ATCCCGGAAGAACTGCGCGCTGCTGCCGAAAGAGTCTCGTTACTGTTACGTGAATTTTG 56335
Qy 427 TTATGTGCCGATGGTGAACCGTGGCTTCCCGGTCTACCGTCTGTTCCCTGTGTCAACGTTA 486
Db 56336 TTATGTGCCGATGGTGAACCGTGGCTTCCCGGTCTACCGTCTGTTCCCTGTGTCAACGTTA 56395
Qy 487 AGCGGGCCGAGCTGGCGCTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTGTTC 546
Db 56396 AGCGGGCCGAGCTGGCGCTTACAAAAATTGGGTAAACGCCGTTTAGGACGCTATCTGTTC 56455
Qy 547 ACATCATCGACATTAAACCGGAGCTTTATTAGATAGCCGCTGATGCCGGCTGTGGGG 606
Db 56456 ACATCATCGACATTAAACCGGAGCTTTATTAGATAGCCGCTGATGCCGGCTGTGGGG 56515
Qy 607 CGAGCTTCCCGCTGCGGATTAAGCGGTAAACCGCTGTTGCTAAAGAACTGTTTTACCG 666
Db 56516 CGAGCTTCCCGCTGCGGATTAAGCGGTAAACCGCTGTTGCTAAAGAACTGTTTTACCG 56575
Qy 667 CGGTACCGTTGTACTAA 684
Db 56576 CGGTACCGTTGTACTAA 56593

RESULT 15

ECOW89

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ECOW89 176195 bp DNA linear BCT 17-DEC-1993
E. coli chromosomal region from 89.2 to 92.8 minutes.

U00006

U00006.1 GI:409785

Escherichia coli

Escherichia coli

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Escherichia.

1 (bases 1 to 176195)

Blattner,F.R., Burland,V., Plunkett,G. III, Sofia,H.J. and

Daniels,D.L.

Analysis of the Escherichia coli genome. IV. DNA sequence of the

region from 89.2 to 92.8 minutes

Nucleic Acids Res. 21 (23), 5408-5417 (1993)

8265357

2 (bases 1 to 176195)

Blattner,F.R.

Direct Submission

Submitted (03-SEP-1993) 608-263-7459

On Oct 29, 1993 this sequence version replaced gi:396288.

This sequence was determined as part of the E. coli Genome Project

(Frederick R. Blattner, director) at the University of

Wisconsin-Madison. Supported by award HG00301 from the NIH Human

Genome Project. The entire sequence was independently determined

from E. coli MGI655; overlaps and conflicts with other sequence

determinations are annotated. The start of this entry overlaps the

end of the entry ECOW87 (U19201) by 93 bp. NOTE: An update was

submitted on 25-OCT-1993, reflecting a correction to the heme

sequence.

Location/Qualifiers

1..176195

/organism="Escherichia coli"

/mol_type="unassigned DNA"

/strain="K-12"

/sub_strain="MG1655"

/db_xref="taxon:562"

/note="This sequence comprises the following lambda

clones: EC18-126, EC14-54, EC21-52, EC17-24, EC17-203,

EC19-105, EC18-208, EC18-110, EC17-200, EC18-33, EC30-262,

EC30-41, EC30K637-2a, EC30K637-5, EC27-975, EC22-27;

M13mp19 or Janus vectors were used for subcloning"

```

misc_feature <1..1737
/note="corresponds to M21516; ECOKATGA(1225..2805)"/
gene 1..1257
/gene="katG"
CDS <1..1257
/gene="katG"
/EC_number="1.11.1.6"
/standard_name="catalase HPI; catalase-peroxide; catalase"
/note="CG Site No. 14983"
/codon_start=1
/transl_table=11
/product="catalase hydroperoxidase I"
/protein_id="AAC43048.1"
/db_xref="GI:396289"
/translation="AITSGLVGVVWVQTPTQWNSYFENLKYEWVQTRSPAGAIQFEA
VPAPEIIPDFPDPKRRKTMVLTDLTRDFEFKISRFLNDPQAFNEAFARXWFK
LTHRDMPKPSRYIGPVKEDILWDPQPIYNTPEQDIIDLKFAIDSGISVSELV
SWAASGTFRGDKRGAGARLALMPORDVNAQAAVRALPVLEKIQKSGKASLA
DIVLGVVGVKAAAGLSHVPAPEVDARODTDIEMFELLEPLADGFRYRA
RDVSTTESLLDKAQQLITAPETALVGGMRVLDGSKNGVFTDRVGVLEVSDF
FVNLDMRYEWKATDESKELFEGRDRETEGVKFTASRADLVFGNSVLVRAVEVYASS
DAHEKFKVDFAAVKVMNLDKDFLL"
misc_feature <1..161
/gene="katG"
/note="corresponds to lambda clone EC18-126"
misc_feature 1..93
/gene="katG"
/note="93 bp overlap with end of L19201 (BCOMU87)"
misc_feature 156..15760
/note="G in M21516; C here"
misc_difference 1266..1305
/note="putative"
terminator 1317..2255
/note="similar to Desulfurolobus ambivalens hypoth. 28.3
kba protein in sor 3' region"
CDS <1..1257
/codon_start=1
/transl_table=11
/protein_id="AAC43049.1"
/db_xref="GI:409786"
/translation="MGERKLHTGSLMSAAGKNPLAIGSLVVLTLTQWNSVFMKQVT
SYIGAFDFALRCIFGALVFLVLLRGRMRPTPKYTLAALTQCMWGLAQWAL
VSGGAKVAILSYTMFPWVVI FAALFLGERLRGQVFAILIAAFLGLVLPQWLDFA
SKMSAMAILSGVSGWASAIKARLVARHPRVDLLSLTSWOMLYAALVMSVVALVPO
REIDQPTVFMAVSAIATALANSLWLFVLKPLPASIASLSTLAVPCVGLFSWL
LGENPGAVGSGGLVILVLAALVSRKKKAAVSKRI"
misc_difference 1429..1432
/note="TTT in M21516; TTTT here"
misc_difference 1720
/note="C in M21516; G here"
CDS complement(2282..2899)
/codon_start=1
/transl_table=11
/label=ORF f205
/protein_id="AAC43050.1"
/db_xref="GI:396291"
/translation="MKASLALLSLTFTSHSLKSPAVPTVVQIQTANTLAIADGAR
QIGSTLFDPAVYVQLTYPGDVQPQGVCSVWVIRALRSQKVDLQKLVHEDAKNFA
EYQKWLKRPSNDIHRVPLETWFSRDKTRPTSKNPSDYQAGDIVSWRLDNGLA
HIGVSDGFARDGTPLVIIHIGAGQEDVLFNWRVMVGHYRFVK"
complement(2886..2915)
promoter <1..1257
/note="promoter matrix score of 42; putative; within ORF
f205, which would suggest alternate start codon"
misc_structure 2912..3053
/note="predicted bend of 81 degrees"
repeat_region 3031..3064
/standard_name="REP; repetitive extragenic palindromic
element"
/note="contains 1 REP sequence"
misc_structure 3090..3219
/note="putative"

```

```

gene <1..1737
/note="predicted bend of 75 degrees"
complement(3174..4316)
/gene="gldA"
CDS complement(3174..4316)
/gene="gldA"
/note="similar to Bacillus stearothermophilus glycerol
dehydrogenase"
/codon_start=1
/transl_table=11
/label=ORF f380
/product="glycerol dehydrogenase"
/protein_id="AAC43051.1"
/db_xref="GI:396292"
/translation="WPHLALLISKGAIMDRITQSPGKYIQGADVINRLGELYLKLPLXER
WLVWDGKFLVGAQTFVKSFKDAGLVVEIAPFGECSQNEIDRLRGIAETACCGAIL
GIGGKLTLDTAKALAHFMGVPAIAPTASTADPCSAISVIYDDEGEFAYLLPNP
NNVIYDTIKVAGAPARLLAAGIGDALATWFEARCSRGATMAGCKTQAALALAEI
CYNTHLEGEKVALEOHVTPPALERTVETANTYLSGVGFSGGLAAAHAVHNGLTAI
PDHLYHGEKVAFGTLTQLVLENAPVEIETVAALSHAVGLPITLAQLDIDKEDVPAK
MRIVAAACAEGETTHNPGGATPQVYAALLVADQYQGRFLQWE"
4022..4156
misc_structure <1..161
/note="predicted bend of 75 degrees"
complement(4288..4950)
/note="similar to Bacillus subtilis hypoth. 20 kba
protein, in tsr 3' region"
CDS <1..161
/transl_table=11
/label=ORF f220
/protein_id="AAC43052.1"
/db_xref="GI:396293"
/translation="MELYLDTANVAEVELARIFPIAGVTTNPSIIAASKESIWEVL
RLQKAIQDEGILFQAQMRDAGMVEEAKRLDAIPGVIVKIPVTSEGLAAIYLKKE
GITTAGTAVYSAAGLLAAGAKVAPYVNRVDAQGGGQIRTVQELQTLLENHAPES
MVLAAAFKTPQALDCLLAGCESITPLDVAQQMLNTPAVESAIEKFEHDMNAAFGTT
HL"
promoter complement(4922..4949)
/note="promoter matrix score of 47; putative; within ORF
f220, which would suggest alternate start codon"
CDS complement(4962..7097)
/note="similar to phosphotransferase system enzyme I"
/codon_start=1
/transl_table=11
/label=ORF f711
/protein_id="AAC43053.1"
/db_xref="GI:409787"
/translation="MCSGSAGGILTPISLNLNALGNLPAKGVDAEQSALENGLTIV
LKNIBFLDSDGATSAILEAHRSLAGDTSLEHLLAGVSAGLSCEAIVASANHFCE
EFSRSSVYQERALDVRDVCFQQLQIYGEORFPAPGLTQPAICMADELTQSFLE
LDKNHLKGLLKSQGTSTHTVILARSFNIPTLVGVVIDALTQWQQTIVIDGNAGAI
VEPGAARYQOEARVQDALREQRVWLTQARTADGIRIRIANNIAHSVEQAARFG
NGAEGVGLFRTMLYMDRTSAPGESELYNIFCOALESANGRSIIVRTMDIGGPKVDY
LNIPAEANPFLGYRAVRIEYASLTQLRSLIRASAHGSLKIMIPMISSEELIWL
KEKLAEQQLRNEHIPDEKTLQGLTLMLEVPVSMFIQDCCBEIDFGSGNDLTQYL
LAVDRNAKVTRYHNSLPAPFLRALDYAVQVHROGKWIIGLCELGAKGSLVLLVGL
GLDELMSAPSIAPAKARMAOLDRECKLLAQAACRTSLEVEHLLAOFRTMQDAP
LVTAECITLESNRSKEEVLKGMTNLLAGRCRPRKLEADLWAEANFSTGLGSFP
AIPHSKSHIEQSTISVARLPQAPVRWGDDEAQFIIMLTUNKHAAGQHMRFPSRLARR
IMHEEFNALVNAASADAISLLQHELEL"
7099..7428
CDS <1..161
/transl_table=11
/label=ORF o109
/protein_id="AAC43054.1"
/db_xref="GI:409788"
/translation="WRARIICGFRILVSETSGSSSSDLTSASGASQWGNSSRSHWLKR
WCACSSSAPEISNWQLSPASVPI SARALLALPCQCAGYATQFVMKYYTAFQRL
PDHVR"
complement(7108..7136)
/note="promoter matrix score of 46; putative"
7200..75839
misc_feature <1..161
/note="corresponds to lambda clone EC21-52"
7578..7699
terminator <1..161
/note="putative"

```

```
CDS
7771..8850
/note="similar to phosphotransferase system enzyme II"
/codon_start=1
/transl_table=11
/label=ORF_0359
/protein_id="AAC43055.1"
/db_xref="GI:396296"
/translation="MNEIVQILKNTROHMTGVSHMIPFVWSGGILLAVSVMLYCKGA
VPDAVDADNLKLPDIFGVAGLTLMPFLAAVIGYSIAERSALAPCAIGARVGNSEFGAG
FPGALIAGLIGIVVHYDKPIPVHVKLSVMPFIPIPIVGLITAGIMMGUGPEFVGA
LTSNLTQMQGQGSIVMLAVIMGLMIAFDMGPNKVAFAFMLICVAOQGVYTVVAI
AAVGICIPPLGLMLATLIGRKNFSAERETKAAALVMGCVGVTEGAI PFAAADPLRVI

Query Match      72.8%; Score 498; DB 15; Length 176195;
Best Local Similarity 100.0%; Pred. No. 3.4e-143; Mismatches 0; Indels 0; Gaps 0;
Matches 498; Conservative 0;

QY 187 ATGTCACACCCCGCGTTAACGCAACTGCGTGCCTGCGCTATTGTAAAGAGATCCCTGCC 246
Db 117748 ATGTCACACCCCGCGTTAACGCAACTGCGTGCCTGCGCTATTGTAAAGAGATCCCTGCC 117807

QY 247 CTGGATCCGCAACTGCTGACTGGCTGTCTGGAGGATTCATCACAAAACGTTTTCGA 306
Db 117808 CTGGATCCGCAACTGCTGACTGGCTGTCTGGAGGATTCATCACAAAACGTTTTCGA 117867

QY 307 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGCGAGCAGAAATGAA 366
Db 117868 CAGCAGGGAAAAACGGTAAGCGTGACGATGATCCGCGAAGGGTTTGTGCGAGCAGAAATGAA 117927

QY 367 ATCCCGCAGAGAACTCCCGTGTCTGCCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTTG 426
Db 117928 ATCCCGCAGAGAACTCCCGTGTCTGCCGAAAGAGTCTCGTTACTGTTACGTGAAAATTTTG 117987

QY 427 TTATGTCCGATGGTGAACCGTGGCTTGCCGGTCTGACCGTGGTTCTCTGTGTCAACGTTA 486
Db 117988 TTATGTCCGATGGTGAACCGTGGCTTGCCGGTCTGACCGTGGTTCTCTGTGTCAACGTTA 118047

QY 487 AGCGGGCCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCCGTTAGGACGCTATCTGTTTC 546
Db 118048 AGCGGGCCGGAGCTGGCGTTACAAAAATTGGGTAAAAACCCGTTAGGACGCTATCTGTTTC 118107

QY 547 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGCCGCTGATGCCGGGCTGTGGGGG 606
Db 118108 ACATCATCGACATTAAACCGGACCTTTATTGAGATAGCCGCTGATGCCGGGCTGTGGGGG 118167

QY 607 CGACGTTCCCGCTCGGATTAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTTACCG 666
Db 118168 CGACGTTCCCGCTCGGATTAGCGGTAAACCGCTGTTGCTAAACAGAACTGTTTTTACCG 118227

QY 667 GCGTCACCGTTGTACTAA 684
Db 118228 GCGTCACCGTTGTACTAA 118245
```

Search completed: July 11, 2006, 20:23:15
Job time : 4397 secs

Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 2
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-718-311-15 (1-62) x AX382258 (1-684)

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTTCTTACGAGCTGTTCACACGACGACGATTTACACAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTCACCTTTCTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 180

Qy 61 Ttrp 62
Db 181 TGGCAT 186

RESULT 5
TOMRBCSB 599 bp mRNA linear PLN 27-APR-1993
LOCUS Tomato (L. esculentum) ribulose-1,5-bisphosphate
DEFINITION carboxylase (RBCS) small subunit mRNA, clone Rbcs-2A.
ACCESSION M13543
VERSION M13543.1 GI:170497
KEYWORDS ribulose 1,5-bisphosphate carboxylase/oxygenase.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum

REFERENCE
AUTHORS Pichersky, E., Tanksley, R., Tanksley, S.D. and Cashmore, A.R.
TITLE Evidence for selection as a mechanism in the concerted evolution of
Lycopersicon esculentum (tomato) genes encoding the small subunit
of ribulose-1,5-bisphosphate carboxylase/oxygenase
Proc. Natl. Acad. Sci. U.S.A. 83 (11), 3880-3884 (1986)
JOURNAL 3012537
PUBMED
COMMENT Original source text: Tomato (L. esculentum) cDNA to mRNA, clone
Rbcs-2A.
Clean copy sequence for [1] kindly provided by E. Pichersky,
31-OCT-1986.

FEATURES
source
1. 599
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
CDS
15..557
/note="ribulose-1,5-bisphosphate carboxylase/ oxygenase
small subunit (EC 4.1.1.39)"
/codon_start=1
/protein_id="AAA34189.1"
/db_xref="GI:170498"
/translation="MASSVISAAVATRSNTQASVAPFTGLKSSATFPVTKQNLD
ITSIASNGRVSCMVQWPPINMKYETLSYLPDLSEQLLESEYLLKNGWVPCLEFE
TEHGFVYRENNKSPGYDGRYTWKLPFMFGCTDTQVLAEOEAKAYPOAWVRIG
FDNRVQVCISFIAYKPEGY"

ORIGIN
Chromosome 3, Rbcs-2 locus.
Alignment Scores:
Pred. No.: 3,99e-28 Length: 599
Score: 297.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.4% Indels: 0
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSB (1-599)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTTCTTACGAGCTGTTCACACGACGACGATTTACACAGCT 60
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 61 AGCATGGTTCACCTTTCTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 120
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 121 CAAACCTTGACATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 180

Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 15 ATGGCTTCCTCTGTCATTTCTTACGAGCTGTTCACACGACGACGATTTACACAGCT 74
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 75 AGCATGGTTCACCTTTCTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 134
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 135 CAAACCTTGACATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 194
Qy 61 Ttrp 61
Db 195 TGG 197

RESULT 6
TOMRBCSE 742 bp mRNA linear PLN 27-APR-1993
LOCUS Tomato RuBP carboxylase small subunit LESS5 mRNA.
DEFINITION M15236
ACCESSION M15236.1 GI:170503
KEYWORDS carboxylase; ribulose bisphosphate carboxylase.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum

REFERENCE
AUTHORS McKnight, T.D., Alexander, D.C., Babcock, M.S. and Simpson, R.B.
TITLE Nucleotide sequence and molecular evolution of two tomato genes
encoding the small subunit of ribulose-1,5-bisphosphate carboxylase
Gene 48 (1), 23-32 (1986)
JOURNAL 3557127
PUBMED
COMMENT Original source text: Tomato (VF36), cDNA to mRNA, clone LESS5.
Draft entry and computer-readable copy of sequence in [1] kindly
provided by T.D. McKnight (15-APR-1987).

FEATURES
source
1. 742
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
mRNA
<1..742
/product="RuBP mRNA"
CDS
28..570
/note="ribulose-1,5-bisphosphate carboxylase, small
subunit precursor"
/codon_start=1
/protein_id="AAA34192.1"
/db_xref="GI:170504"
/translation="MASSVISAAVATRSNTQASVAPFTGLKSSATFPVTKQNLD
ITSIASNGRVSCMVQWPPINMKYETLSYLPDLSEQLLESEYLLKNGWVPCLEFE
TEHGFVYRENNKSPGYDGRYTWKLPFMFGCTDTQVLAEOEAKAYPOAWVRIG
FDNRVQVCISFIAYKPEGY"

sig_peptide
mat_peptide
ORIGIN 614 bp upstream of DraI site.
Alignment Scores:
Pred. No.: 5,27e-28 Length: 742
Score: 297.00 Matches: 61
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 97.4% Indels: 0
DB: 4 Gaps: 0

US-10-718-311-15 (1-62) x TOMRBCSE (1-742)
Qy 1 MetAlaSerSerValIleSerSerAlaAlaValAlaThrArgSerAsnValThrGlnAla 20
Db 1 ATGGCTTCCTCTGTCATTTCTTACGAGCTGTTCACACGACGACGATTTACACAGCT 74
Qy 21 SerMetValAlaProPheThrGlyLeuLysSerSerAlaThrPheProValThrLysLys 40
Db 75 AGCATGGTTCACCTTTCTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 134
Qy 41 GlnAsnLeuAspIleThrSerIleAlaSerAsnGlyGlyArgValSerCysMetGlnVal 60
Db 135 CAAACCTTGACATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTT 194
Qy 61 Ttrp 61
Db 195 TGG 197

THIS PAGE BLANK (USPTO)